Journal of Marketing

CUK-H05815-13-KP2610 .

January 2006
Volume 70, Number 1

A Quarterly Publication of the
American Marketing Association

From the Editor 1 Roland T. Rust

Customer Satisfaction and Stock Prices: High Returns, Low Risk 3 Claes Fornell, Sunil Mithas, Forrest V. Morgeson III, & M.S. Krishnan

Market Pioneer and Early Follower Survival Risks: A Contingency Analysis of Really New Versus Incrementally New Product-Markets 15 Sungwook Min, Manohar U. Kalwani, & William T. Robinson

Brands as Signals: A Cross-Country Validation Study 34 Tülin Erdem, Joffre Swait, & Ana Valenzuela

Emotional Branding and the Strategic Value of the Doppelgänger Brand Image 50 Craig J. Thompson, Aric Rindfleisch, & Zeynep Arsel

Does the Internet Promote Better Consumer Decisions? The Case of Name-Your-Own-Price Auctions 65 Martin Spann & Gerard J. Tellis

Understanding Consumer Responses to Product Risk Information 79 Anthony D. Cox, Dena Cox, & Gregory Zimet

How Does Enforcement Deter Gray Market Incidence? 92 Kersi D. Antia, Mark E. Bergen, Shantanu Dutta, & Robert J. Fisher

When Should a Retailer Create an Exciting Store Environment? 107 Velitchka D. Kaltcheva & Barton A. Weitz

Value-Based Differentiation in Business Relationships: Gaining and Sustaining Key Supplier Status 119 Wolfgang Ulaga & Andreas Eggert

Job Satisfaction, Job Performance, and Effort: A Reexamination Using Agency Theory 137 Markus Christen, Ganesh Iyer, & David Soberman

13

AMERICAN MARKETING ASSOCIATION

"Covers brand management, brand extension, name changes, brand revitalization, and global brand management. Recommended." —CHOICE magazine

"Creatively written…challenging in its thinking yet interesting and accessible." —Media Week

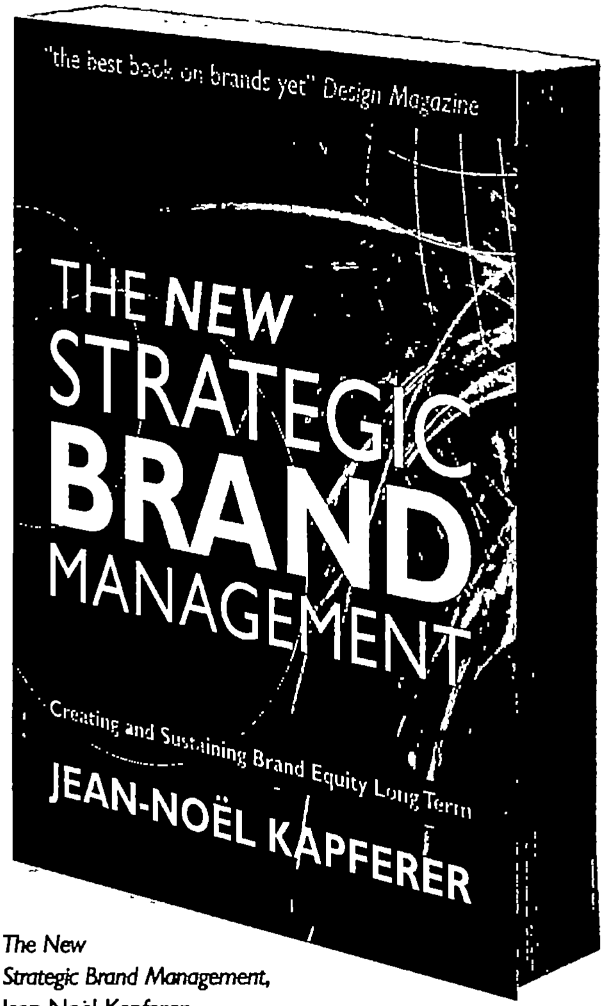
THE MOST INNOVATIVE, THOROUGH AND RELEVANT BOOK ON BRAND MANAGEMENT

Innovative: proposes new models, new concepts, new benchmarks for brand and business building, all absent from other books on brands. New approaches to brand identity and positioning, brand architectures, brand portfolios and brand globalization.

Thorough: analyzes in depth all facets of brand management, with practical models and advice.

Relevant: no other book addresses the key issues of brand management today including: **Building brands in saturated markets • What innovation for what brands? • Managing retail brands • Activating the brand • Energizing the brand value chain • Globalizing existing brands • Brand extensions.**

KP2610


"the best book on brands yet" Design Magazine
THE NEW STRATEGIC BRAND MANAGEMENT
Creating and Sustaining Brand Equity Long Term
JEAN-NOËL KAPFERER

The New Strategic Brand Management,
Jean-Noël Kapferer
Paper, 0 7494 4283 2, $55.00
Published by Kogan Page

About the author: **JN Kapferer** (HEC, Ph.D Northwestern University) is considered one of the top three branding experts in the world. His style of analysis uniquely blends key research developments with very practical concerns. He has experience as a consultant for European, Asian and American companies and teaches at advanced executive development seminars worldwide.

KOGAN PAGE

For more details, full table of contents, to download a sample chapter and order, visit
www.kogan-page.co.uk/kapferer or
www.styluspub.com
Inspection copies available

Journal of Marketing

Vol. 70, No. 1

AMERICAN MARKETING ASSOCIATION

EDITOR
Roland T. Rust
Robert H. Smith School of Business
3451 Van Munching Hall
University of Maryland
College Park, MD 20742-1815

ASSISTANT TO THE EDITOR
Marti Tiedeman
E-mail: jom@ama.org
Tel: (312) 542-9024

Millicent Locke
University of Maryland

WEB EDITOR
Siva K. Balasubramanian
Southern Illinois University

MARKETING LITERATURE REVIEW
Myron Leonard
Western Carolina University

BOOK REVIEWS
Ramarao Desiraju
University of Central Florida

VICE PRESIDENT—PUBLICATIONS
Russell S. Winer
New York University

CHIEF EXECUTIVE OFFICER
Dennis Dunlap

GROUP PUBLISHER
Jack Hollfelder

EDITORIAL STAFF
Phone: (312) 542-9000
Fax: (312) 542-9001

MANAGING EDITOR
Francesca Van Gorp Cooley

TECHNICAL EDITOR
Andy Seagram

GRAPHICS EDITOR
Mary Loye

Advertising Staff
Fax: (312) 922-3763

Richard Ballschmiede, Advertising
(312) 542-9076

Sally Schmitz, Production
(312) 542-9038

JM Web site:
Submissions: www.marketingjournals.org/JM
Main page: www.marketingpower.com/JM

The *Journal of Marketing* (ISSN 0022-2429) is published quarterly (January, April, July, and October). *Annual Subscription Rates*: Subscriptions are $90 for individuals in the United States and Canada and $115 in other countries. Library, school, and business subscriptions are $225 in the United States and Canada and $250 in other countries. Subscriptions for AMA members are $53. Prices are for one year. Single copy rate. $50 individual, $50 corporate in the United States and Canada; add $5 in other countries. Prepayment is required in U.S. dollars or the equivalent

Published by the Publications Group of the American Marketing Association, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266 Periodicals postage paid at Chicago, IL and additional mailing offices. POSTMASTER: Send address changes to the *Journal of Marketing*, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266, USA.

•**MANUSCRIPTS**. Electronic submissions must be made via the *JM* Web site at http://www.marketingjournals org. Paper submissions are not accepted Editorial policies, review board members, Tables of Contents for *JM*, and so on, are available on the *JM* Web site at http://www.marketingpower.com/jm.

•**ADVERTISERS AND ADVERTISING AGENCIES** assume liability for all content (including text, representations, and illustrations) of advertisements published, and also assume responsibility for any claims arising therefrom made against the publisher The right is reserved to reject any advertisement.

•**SUBSCRIPTIONS***:* Members: Send to American Marketing Association, 311 S. Wacker Dr, Suite 5800, Chicago, IL 60606-2266 (1-800-AMA-1150 or 312/542-9000), or e-mail to info@ama.org Nonmembers Send to American Marketing Association, P.O. Box 11806, Birmingham, AL 35202, 1-800-633-4931, 205-995-1567, Fax. 205-995-1588, e-mail: amasubs@ebsco.com. Canadian residents add 7% GST (GST registration # 12747852).

•**CHANGE OF ADDRESS:** Send new address, address label from *JM*, and zip code to American Marketing Association. Allow 6 to 8 weeks.

•**PERMISSIONS AND REPRINT INFORMATION***:* Where necessary, permission is granted by the copyright owner for those registered with the Copyright Clearance Center (CCC), 222 Rosewood Drive, Danvers, MA 01923, to photocopy any article herein for a flat fee. Contact the CCC with the ISSN (0022-2429), volume, and first and last page numbers of each article copied Copying for other than personal or internal reference without express permission of the American Marketing Association Publications Group is prohibited Address requests for customized bulk reprints to Kristen Ramsey, Valeo IP, 651-582-3851 or 1-800-217-7874

The views expressed in *JM* are not necessarily those of the editorial staff or the publisher. **All articles in this journal, unless otherwise noted, have undergone a blind review process.**

Copyright © 2006, American Marketing Association All rights reserved.

Editorial Review Board

ROHINI AHLUWALIA
University of Minnesota

KUSUM L. AILAWADI
Dartmouth College

TIM AMBLER
London Business School

EUGENE W. ANDERSON
University of Michigan

TOR WALLIN ANDREASSEN
Norwegian School of Management

ASIM ANSARI
Columbia University

ERIC J. ARNOULD
University of Arizona

BARRY L. BAYUS
University of North Carolina

PAUL D. BERGER
Boston University

PAUL N. BLOOM
University of North Carolina

LISA E. BOLTON
University of Pennsylvania

RUTH N. BOLTON
Arizona State University

ROBERT BORDLEY
General Motors Corporation

DOUGLAS BOWMAN
Emory University

RODERICK J. BRODIE
University of Auckland

DIPANKAR CHAKRAVARTI
University of Colorado

PIERRE CHANDON
INSEAD

PRADEEP K. CHINTAGUNTA
University of Chicago

KEVIN CLANCY
Copernicus Marketing Consulting

PETER J. DANAHER
University of Auckland

JOHN A. DEIGHTON
Harvard University

MARNIK G. DEKIMPE
Catholic University Leuven

WAYNE S. DESARBO
Pennsylvania State University

RAVI DHAR
Yale University

NAVEEN DONTHU
Georgia State University

TULIN ERDEM
University of California, Berkeley

PETER S. FADER
University of Pennsylvania

GARY L. FRAZIER
University of Southern California

PETER N. GOLDER
New York University

RAJDEEP GREWAL
Pennsylvania State University

ABBIE GRIFFIN
University of Illinois, Urbana-Champaign

SUNIL GUPTA
Columbia University

JOHN R. HAUSER
Massachusetts Institute of Technology

JAN B. HEIDE
University of Wisconsin, Madison

CHRISTIAN HOMBURG
University of Mannheim

WAYNE D. HOYER
University of Texas at Austin

G. TOMAS M. HULT
Michigan State University

J. JEFFREY INMAN
University of Pittsburgh

GANESH IYER
University of California, Berkeley

CHRIS JANISZEWSKI
University of Florida

SANDY D. JAP
Emory University

BERNARD J. JAWORSKI
Monitor Group

MICHAEL D. JOHNSON
University of Michigan

P.K. KANNAN
University of Maryland

TIMOTHY L. KEININGHAM
Ipsos Loyalty

AMNA KIRMANI
University of Maryland

AJAY K. KOHLI
Emory University

ROBERT V. KOZINETS
York University

V. KUMAR
University of Connecticut

DONALD R. LEHMANN
Columbia University

KATHERINE N. LEMON
Boston College

MARY FRANCES LUCE
Duke University

ROBERT F. LUSCH
University of Arizona

JOHN G. LYNCH JR.
Duke University

DEBORAH J. MACINNIS
University of Southern California

VIJAY MAHAJAN
University of Texas at Austin

CARL F. MELA
Duke University

VIKAS MITTAL
University of Pittsburgh

CHRISTINE MOORMAN
Duke University

PAGE MOREAU
University of Colorado

SCOTT A. NESLIN
Dartmouth College

STEPHEN M. NOWLIS
Arizona State University

JOSEPH C. NUNES
University of Southern California

A. PARASURAMAN
University of Miami

WILLIAM D. PERREAULT
University of North Carolina

F.G.M. PIETERS
Tilburg University

JAIDEEP PRABHU
Imperial College

ARVIND RANGASWAMY
Pennsylvania State University

BRIAN T. RATCHFORD
University of Texas at Dallas

AMERICUS REED
University of Pennsylvania

WERNER REINARTZ
INSEAD

STEVEN M. SHUGAN
University of Florida

ITAMAR SIMONSON
Stanford University

DILIP SOMAN
University of Toronto

JOYDEEP SRIVASTAVA
University of Maryland

JAN-BENEDICT E.M. STEENKAMP
Tilburg University

DAVID W. STEWART
University of Southern California

STEFAN STREMERSCH
Erasmus University Rotterdam

BAOHUNG SUN
Carnegie Mellon University

GERARD J. TELLIS
University of Southern California

STIJN VAN OSSELAER
Erasmus University Rotterdam

P. RAJAN VARADARAJAN
Texas A&M University

STEPHEN L. VARGO
University of Hawaii, Manoa

PETER VERHOEF
University of Groningen

MICHEL WEDEL
University of Michigan

BARTON A. WEITZ
University of Florida

RUSSELL S. WINER
New York University

ANTHONY ZAHORIK
Burke Institute

VALARIE A. ZEITHAML
University of North Carolina, Chapel Hill

Z. JOHN ZHANG
University of Pennsylvania

KP 2610

From the Editor

It is a great honor to assume the editorship of *Journal of Marketing* (*JM*). My term of editorship began on July 2005 and continues through June 2008, though the first issue for which I am responsible is January 2006. The articles from the July 2005 and October 2005 issues were all accepted by the previous editor, Ruth Bolton, who has done a masterful job in building on the illustrious editorial tradition that had been established over the years by her predecessors. I am very grateful to Ruth and to all the previous editors for helping make *JM* the journal it is today.

Current Status

The oldest of the four major journals in marketing, *JM* enjoys a prestigious position as the best-cited marketing journal (in two of the last three Institute for Scientific Information citation impact studies), the most selective marketing journal (acceptance rate of 11%), and the marketing journal that receives the most submitted manuscripts (345 in calendar year 2004). In addition, *JM* has ranked first (Baumgartner and Pieters 2003), second (Theoharakis and Hirst 2002), and first (Hult, Neese, and Bashaw 1997) in the three most recent surveys of the leading marketing journals. It is clear that *JM* enjoys an enviable position among marketing's most elite academic journals. My charge as editor is to do my utmost to maintain and, hopefully, to enhance the journal's premier status.

The Positioning of *JM*

Traditionally, *JM* has been positioned as the broadest, the most readable, and the closest to management practice of the four leading marketing journals. Its two greatest strengths are typified by its two most long-standing article awards: the Harold H. Maynard Award for marketing theory and the Marketing Science Institute/H. Paul Root Award (formerly the Alpha Kappa Psi Award) for advancing the practice of marketing. Substantively, *JM* spans the entire field of marketing, so its domain overlaps considerably with that of the other major marketing journals.

It is important that *JM* maintains its historical strengths. These include publishing the best articles that advance marketing theory and the best articles that are relevant to marketing practice. During my editorship, *JM* will reinforce its historical strengths while maintaining as broad a positioning as possible. Essentially, *JM* attempts to publish the best substantive articles in marketing. As such, the best articles that are currently published in *Journal of Marketing Research* and *Marketing Science* that are not primarily focused on methodology would be welcome in *JM*. Likewise, the best *Journal of Consumer Research* articles that are relevant to marketing would also be welcome in *JM*. Ultimately, *JM* will continue to be substantive and problem driven. It will strive to publish the best work in marketing of any form or style as long as it has relevance to marketing theory or marketing practice.

This positioning implies that *JM* will not compromise with respect to sophistication. Authors will be encouraged to use the simplest methods necessary to address their subject, but if more sophisticated methods are required, such an approach should not preclude publication. The use of sophisticated methods for their own sake will not be encouraged, but theoretical or methodological sophistication will be encouraged when they help address a substantive topic.

Breadth of Coverage

I believe that all articles in *JM* should have something substantive to say to one of the primary stakeholders of marketing (marketing managers, consumers, or public policy makers). Aside from this requirement, *JM* will be open to any topic and any research approach. I especially want to emphasize that *JM* will be receptive to all methodological approaches, including (but not limited to) survey research, experimental designs, analytical models, ethnographic research, structural equation models, econometric analysis, and conceptual analysis. All research approaches have strengths and weaknesses, and each submitted paper will be evaluated on its unique merits.

Publishing Truly Important Articles

I have found in my own research that, in general, my most innovative articles have had the toughest time in the review process. Those same articles, once published, have had the greatest impact. As editor, I will go out of my way to be sympathetic to truly innovative papers. I will give such papers extra consideration. On the margin, I would rather accept an innovative paper that is controversial than an incremental paper that is perfectly done but unsurprising. I will give the importance of the problem addressed considerable weight in the publication decision. To achieve this, on occasion, I will be willing to overrule a consensus negative evaluation by the reviewers to give an innovative paper an extra chance. I strongly encourage authors to submit work that has an opportunity to shake things up and change the thinking in the field.

An International Journal

The economy is global, and marketing talent is spread around the world. During the seven years that I served as editor of *Journal of Service Research*, more than half of the papers that I accepted for publication had at least one non-U.S. author. Likewise, non-U.S. authors will have a fair opportunity to publish in *JM*. That is not to say that publication will be easy—recall that the journal's current acceptance rate is 11%—but I will guarantee that there will be no bias against non-U.S. authors or non-U.S. data. I have also increased non-U.S. participation on the Editorial Review Board.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Connecting to the Business Community

Traditionally, *JM* has been the leading outlet for managerial research among the leading marketing journals, and I wish to maintain *JM*'s strength in this area. I also want to encourage the submission of articles by practitioners. Because the standard of research sophistication at *JM* is high, coauthorship with academics is often the easiest way for a practitioner to create a publishable article, though some practitioner authors with a prior record of publication in academic journals may well be able to produce a publishable article on their own.

The Editorial Review Board

I have appointed an Editorial Review Board of 90 experts, a slightly smaller board than in recent years. Most of the members of the Editorial Review Board are holdovers from Ruth Bolton's board. To attempt to strengthen and broaden the board, I have altered its composition in a few important ways. First, I have added new board members from a cross-section of the marketing field and have particularly emphasized adding people with a broad point of view and an active, recent research record. Second, I have increased the percentage of non-U.S. board members to the highest percentage at any of the four leading marketing journals. Third, to increase the connection to the business world, I have appointed five new members from the business community, all of whom have significant publication experience.

The Review Process

I will assign two to four reviewers to each manuscript and attempt to have the reviewers represent different points of view. At *Journal of Service Research*, I maintained an average turnaround time for new manuscripts of approximately 40 days, which I believe was the fastest journal turnaround of any marketing-related journal. I will work hard to maintain a culture of prompt reviews at *JM* as well. If one of the assigned reviewers is slow, I may sometimes proceed to a decision based on two reviewers rather than wait for all reviews to arrive. I will replace habitually tardy review board members. With most papers, I intend to make a decision in two rounds of review, but the most innovative papers may require more rounds because, by definition, such papers break new ground and thus tend to be more difficult.

Conflicts of Interest

To address potential conflicts of interest, I am formalizing some policies. In most cases, these policies have already been informally maintained at *JM* by previous editors. First, a guest editor (chosen from the former editors of *JM*) will make decisions about any manuscripts submitted by the editor's current faculty colleagues, current or former doctoral students, and current business associates. Second, for all award voting conducted by the Editorial Review Board, voters will be prohibited from voting for their own articles or for article by colleagues from the same university or company, and the editor's articles will not be eligible for any award during the editor's term if it is selected by the Editorial Review Board that the editor has appointed.

Administration of the Journal

The American Marketing Association (AMA) is instituting a new process for administering the journal. Under this process, the editorial assistant for the journal, Marti Tiedeman, is located at the AMA offices in Chicago rather than at the editor's university. In addition, I have hired a local editorial assistant, Millicent Locke, for the College Park office. Because all communication is now electronic and conducted through the Internet, this new procedure is anticipated to have minimal negative impact. In fact, the editor and editorial assistant have been in separate locations ever since the previous editor, Ruth Bolton, moved to Arizona State University.

Strengthening the *JM* Web Site

The AMA is adding value to its Web site overall, and the same will be done with the *JM* Web site, which is part of the AMA site. Because of the added emphasis on the Web site, I have appointed a Web site editor, Siva Balasubramanian, to plan, manage, and edit the site. Value will be added to the site by migrating several features that previously appeared in the hard copy of the journal. For example, the "Book Review" section, which will be edited by the new book review editor, Ramarao Desiraju, will now appear on the Web site along with each article's summary and author biography and the Marketing Literature Review. Additional features will be added over time.

Summary

Journal of Marketing has a proud history and reputation, and I wish to maintain the journal's strengths and work to make an already fine journal even better. During my editorship, *JM* will maintain its broad positioning and will welcome research of all styles and topics, as long as they are pertinent to marketing theory or marketing practice. My goal is to make *JM* every author's first-choice outlet for substantive research in marketing.

ROLAND T. RUST
Editor

REFERENCES

Baumgartner, Hans and Rik Pieters (2003), "The Structural Influence of Marketing Journals: A Citation Analysis of the Discipline and Its Subareas Over Time," *Journal of Marketing*, 67 (April), 123–39.

Hult, G. Tomas M., William T. Neese, and R. Edward Bashaw (1997), "Faculty Perceptions of Marketing Journals," *Journal of Marketing Education*, 19 (1), 37–52.

Theoharakis, Vasilis and Andrew Hirst (2002), "Perceptual Differences of Marketing Journals: A Worldwide Perspective," *Marketing Letters*, 13 (4), 389–402.

Claes Fornell, Sunil Mithas, Forrest V. Morgeson III, & M.S. Krishnan

Customer Satisfaction and Stock Prices: High Returns, Low Risk

Do investments in customer satisfaction lead to excess returns? If so, are these returns associated with higher stock market risk? The empirical evidence presented in this article suggests that the answer to the first question is yes, but equally remarkable, the answer to the second question is no, suggesting that satisfied customers are economic assets with high returns/low risk. Although these results demonstrate stock market imperfections with respect to the time it takes for share prices to adjust, they are consistent with previous studies in marketing in that a firm's satisfied customers are likely to improve both the level and the stability of net cash flows. The implication, implausible as it may seem in other contexts, is high return/low risk. Specifically, the authors find that customer satisfaction, as measured by the American Customer Satisfaction Index (ACSI), is significantly related to market value of equity. Yet news about ACSI results does not move share prices. This apparent inconsistency is the catalyst for examining whether excess stock returns might be generated as a result. The authors present two stock portfolios: The first is a paper portfolio that is back tested, and the second is an actual case. At low systematic risk, both outperform the market by considerable margins. In other words, it is possible to beat the market consistently by investing in firms that do well on the ACSI.

Does the satisfaction of a company's customers have anything to do with the company's stock price? In a way, it must. Otherwise, it would be necessary to go back to the drawing board of how capitalistic market theory is supposed to work. The empirical literature on the nature of the relationship is growing, but it is still in its infancy in many respects. Empirical results (Anderson 1996; Anderson, Fornell, and Lehmann 1994; Anderson, Fornell, and Mazvancheryl 2004; Bolton 1998; Fornell 2001; Ittner and Larcker 1996, 1998; Rust, Moorman, and Dickson 2002) point toward a significant relationship between customer satisfaction and economic performance in general, but less is known about how the satisfaction of companies' customers translates into securities pricing and investment returns, and virtually nothing is known about the associated risks. The tacit link between buyer utility and the allocation of investment capital is a fundamental principle on which the economic system of free market capitalism rests. The degree to which capital flows from investors actually move in tandem with consumer utility is a matter of significant importance because it is an indication of how well (or poorly) markets truly work. A parallel, but narrower, issue is the extent to which equity markets themselves are efficient. The assertion in efficient markets theory is that financial prices accurately reflect all public information at all times, but financial assets can be correctly priced in equity markets even if consumer sovereignty is restricted in product markets. In a monopoly, for example, dissatisfied buyers cannot punish the seller by taking their business elsewhere, and capital flows are not likely to correspond with consumer utility, but that does not in itself imply that capital markets are inefficient. However, efficient allocation of resources in the overall economy and consumer sovereignty depend on the joint ability of product and capital markets to reward and punish companies such that firms that fail to satisfy customers are doubly punished by both customer defection and capital withdrawal. Similarly, firms that do well by their customers would be doubly rewarded by more business from customers and more capital from investors.

We begin by estimating the relationship between customer satisfaction and market value of equity. As expected, we find a strong relationship. This prompts the question, How do investors react to news about changes in customer satisfaction? Despite the significant relationship between customer satisfaction and market value, we find no evidence that they react in a timely manner. Is this because the news is somehow factored into share prices already, or is it due to market imperfections? The evidence suggests the latter. Two stock portfolios, one paper strategy that is back tested and another describing an actual case, lead to similar conclusions: It is possible to beat the market consistently with investment decisions based on customer satisfaction. Perhaps even more remarkable is that these results are not associated with a risk premium. Indeed, the opposite is true;

Claes Fornell is Donald C. Cook Professor of Business Administration and Director of the National Quality Research Center (e-mail: cfornell@umich.edu), Forrest V. Morgeson III is a research scientist, National Quality Research Center (e-mail: morgeson@umich.edu), and M.S. Krishnan is Professor of Business Information Technology, Area Chair, and Mary and Mike Hallman Fellow (e-mail: mskrish@umich.edu), Stephen M. Ross School of Business, University of Michigan. Sunil Mithas is an assistant professor, R.H. Smith School of Business, University of Maryland (e-mail: smithas@umd.edu). The authors thank John Cattler, Stefan Dragolov, Janice Kwan, Eli Dragolov, and Steven Erlsch for providing research assistance in collecting necessary data for this study. They also acknowledge useful comments and help from David VanAmburg and Jaesung Cha of National Quality Research Center at University of Michigan during the progress of the study. The authors are grateful to Ramanath Subramanyam for his help and comments during early stages of this research. Financial support for this study was provided in part by the Michael R. and Mary Kay Hallman Fellowship at the Stephen M. Ross School of Business, University of Michigan.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

the systematic risk is low relative to the market. Although these results may appear too good to be true and puzzling from a conventional financial perspective, we present arguments suggesting that they are in line with empirical findings both in marketing and in the literature on customer satisfaction.

Customer Satisfaction: Economic Effects

We now turn to the fundamentals behind the relationship between customer satisfaction and stock prices. Both basic neoclassical economics and marketing theory provide a general case for a positive association. Conversely, several market factors may contribute to a negative relationship. In addition to market factors, there are other circumstances in which the association could actually be negative. We discuss both categories and empirically analyze stock market reaction to quarterly news releases from the American Customer Satisfaction Index (ACSI) (Fornell et al. 1996). By controlling for other firm-specific news, we test the effect of news announcements about customer satisfaction on stock prices.

Real economic growth depends on the productivity of economic resources and the quality of the output (as experienced by the user) that those resources generate. In the final analysis, expanding economic activity per se is not what is essential. Both marketing and neoclassical economics view consumer utility, or satisfaction, as the real standard for economic growth. The extent to which buyers financially reward sellers that satisfy them and punish those that do not and the degree to which investment capital reinforces the power of the consumer are fundamental to how markets function. A well-functioning market allocates resources, including capital, to create the greatest possible consumer satisfaction as efficiently as possible. A dissatisfied buyer will not remain a customer unless there is nowhere else to go or it is too expensive to get there.

Restrictions on consumer choice may be good for the monopolist but, in general, are considered harmful to the economy. In a competitive marketplace that offers meaningful consumer choice alternatives, firms that do well by their customers are rewarded by repeat business, lower price elasticity, higher reservation prices, more cross-selling opportunities, greater marketing efficiency, and a host of other things that usually lead to earnings growth (Fornell et al. 1996).

To what degree do markets actually reflect the normative tenets of market theory? Under what circumstances would news about increasing customer satisfaction contribute to higher stock prices? Under what circumstances would it have the opposite effect? From a cursory glance at the literature, it is tempting to assume that news about rising customer satisfaction would have an immediate and positive effect on stock prices. Beginning with the early contributions of Bursk (1966) and Jackson (1985), there is substantial conceptual logic and empirical evidence to suggest that the health of a firm's customer relationships is a relevant indicator of firm performance (Ambler et al. 2002; Bell et al. 2002; Berger et al. 2002; Blattberg and Deighton 1996; Bolton, Lemon, and Verhoef 2004; Fornell 1995, 2001; Fornell et al. 1996; Hogan, Lemon, and Rust 2002; Rust et al. 2004). For example, it has been found that customer satisfaction has a negative impact on customer complaints and a positive impact on customer loyalty and usage behavior (Bolton 1998; Fornell 1992). Increased customer loyalty may increase usage levels (Bolton, Kannan, and Bramlett 2000), secure future revenues (Rust and Keiningham 1994; Rust, Moorman, and Dickson 2002), reduce the cost of future transactions (Reichheld and Sasser 1990), lower price elasticity (Anderson 1996), and minimize the likelihood of customer defection (Anderson and Sullivan 1993; Mithas, Jones, and Mitchell 2004). Customer satisfaction may also reduce costs related to warranties, complaints, defective goods, and field service costs (Anderson, Fornell, and Lehmann 1994; Fornell 1992; Garvin 1988). Empirical evidence also suggests that customer perceptions of superior quality are associated with higher economic returns (Aaker and Jacobson 1994; Capon, Farley, and Hoenig 1990; Fornell 2001). Naumann and Hoisington (2001) report positive associations among employee satisfaction, customer satisfaction, market share, and productivity measures at IBM Rochester. Several case-based research studies also find that customer satisfaction is positively associated with employee loyalty, cost competitiveness, profitable performance, and long-term growth (Heskett, Sasser, and Schlesinger 1997; Reichheld and Teal 1996). Finally, in a recent study of the relationship between customer satisfaction and shareholder return, Anderson, Fornell, and Mazvancheryl (2004) find a strong relationship between customer satisfaction and Tobin's q (as a measure of shareholder value) after controlling for fixed, random, and unobservable factors.

If, as the empirical evidence suggests, customer satisfaction tends to improve repeat business, usage levels, future revenues, positive word of mouth, reservation prices, market share, productivity, cross-buying, cost competitiveness, and long-term growth and if it tends to reduce customer complaints, transaction costs, price elasticity, warranty costs, field service costs, defective goods, customer defection, and employee turnover, it seems logical to expect that these effects will eventually affect stock prices and company valuations. Even if a substantial portion of the empirical findings were dismissed,[1] it would be difficult not to take seriously the notion of customer satisfaction as a real, albeit intangible, economic asset. Putting the argument in the context of economic assets allows for a logical link to conventional capital asset pricing (see Fornell and Wernerfelt's [1987, 1988] economic analysis of the value of keeping customers and the more recent work by Wayland and

---

[1] For example, in a study of a catalog retailer, Reinartz and Kumar (2000, 2002) find a weaker-than-expected (but significant) relationship between customer retention and profitability. The strength of any customer retention–profitability relationship depends on the cost of creating and maintaining repeat customers. If repeat business is created through price discounts or other means that do not cause an upward shift in the firm's demand curve, the relationship will be weaker. Thus, repeat business produced by higher customer satisfaction will be more profitable in general than repeat business generated by price discounts.

Cole [1997], Srivastava, Shervani, and Fahey [1998], Gruca and Rego [2005], Gupta, Lehmann, and Stuart [2004], Anderson, Fornell, and Mazvancheryl [2004], Hogan, Lemon, and Rust [2002], Venkatesan and Kumar [2004], and Rust, Lemon, and Narayandas [2005]). Srivastava, Shervani, and Fahey (1998) provide the conceptual logic for more categorically drawing the link from the empirical findings on customer satisfaction to stock returns and shareholder value. They identify four major determinants of a company's market value: (1) acceleration of cash flows, (2) increase in cash flows, (3) reduction of risk associated with cash flows, and (4) increase in the residual value of the business. Not only does the customer asset, as defined by the expected future discounted net cash flow from current and future customers, influence all four determinants, but it can also be expressed as the sum of all other economic assets of the firm (Wayland and Cole 1997); Gupta, Lehmann, and Stuart (2004) also demonstrate this.

First, the acceleration of cash flows is affected by the speed of buyer response to marketing efforts. Because it takes less effort to persuade a satisfied customer (Keller 1993) and because it is more likely that receivables turnover is better for firms with satisfied customers, speed of cash flow is positively affected. Second, there is also a relationship between customer satisfaction and levels of net cash flows. Marginal costs of sales and marketing are lower, as are the requirements for working capital and fixed investments (Srivastava, Shervani, and Fahey 1998). In addition, revenue growth benefits from more repeat business. Gruca and Rego (2005) provide empirical evidence for this. They show that increases in customer satisfaction lead to significant cash flow growth. Specifically, one point in customer satisfaction (measured by ACSI on a 0–100 scale) was associated with a 7% increase in cash flow. Third, Gruca and Rego also find that the risk associated with future cash flows is reduced for firms with high customer satisfaction. If the variability in cash flows is reduced, the cost of capital goes down as well, thus producing yet another source for stock price growth. Finally, the residual value of business is a function of the size, loyalty, and quality of the customer base, all of which are obviously related to the satisfaction of this customer base (Srivastava, Shervani, and Fahey 1998).

Customer Satisfaction and Market Value of Equity

In light of the evidence reported in the literature, it might be expected that a significant relationship exists between market value of equity and ACSI. If customer satisfaction truly is an economic asset that is not only left off the balance sheet but also not fully reflected by the recorded assets, there should be a relationship between ACSI scores and market capitalization. Following standard practice (Barth and McNichols 1994; Ittner and Larcker 1998; Landsman 1986), the model we define in Equation 1 estimates the effect of an identifiable, but intangible, asset on market capitalization in a combined longitudinal cross-sectional setting, while controlling for accounting book values. In this case, we are interested in (1) whether customer satisfaction (ACSI) has a significant effect on market value of equity and (2) the size of this effect when we control for the influence of the recorded assets and liabilities. Thus:

$$(1) \qquad \ln MVE = \alpha + \beta_1 \ln BVA + \beta_2 \ln BVL + \beta_3 \ln ACSI,$$

where MVE is the market value of equity, BVA is the book value of total assets, BVL is the book value of liabilities, and ACSI is the respective company score on the American Customer Satisfaction Index.

We obtained the data from 1994–2002 ACSI scores and COMPUSTAT company book values for the same years; there were a total of 601 observations. The mean value for market value is $26 billion. For book value of assets and liabilities, the respective means are $23 billion and $15 billion.

Table 1 reports the results of fitting Equation 1 by ordinary least squares. The regression is highly significant with an R-square of .70. The ACSI coefficient is significant and large. A 1% change in ACSI is associated with a 4.6% change in market value. The coefficients for assets and liabilities are also significant with the expected signs. Table 1 also includes estimates without the ACSI variable: R-square drops to .61, and the negative coefficient for liabilities increases. Thus, by omitting customer satisfaction in the estimation of company market value, the negative impact of liabilities appears to be greater than it should be. This is also consistent with the previous discussion that customer satisfaction alleviates cash flow volatility and reduces the cost of capital.

Investor Reaction

We offer the preceding results as empirical evidence in favor of the argument that customer satisfaction should be considered an economic asset beyond what is recognized on the balance sheet. The results are also consistent with the finding that customer satisfaction is associated with long-term shareholder value. However, they do not address

TABLE 1
The Effect of ACSI on Market Value of Equity

		Ln(Market Value)	**Ln(Market Value)**
ACSI (lnACSI)	β_1	4.592 (.000)	
Total assets (lnBVA)	β_2	1.943 (.000)	2.039 (.000)
Total liabilities (lnBVL)	β_3	–1.020 (.003)	–1.157 (.003)
Constant	α	–20.056 (.000)	.235 (.704)
N		601	601
R^2		.70	.61
F		(.000)	(.000)

Notes: We estimated the models by including dummy variables (not shown) for years 1995 through 2002 (1994 was the reference year); *p* values are in parentheses.

whether ACSI scores provide information to investors (that is not already known to them). How then do investors react to news about customer satisfaction? Do they take advantage of customer satisfaction as a leading indicator of financial performance, or do they fail to realize the potential of the information? Is the information already factored in share prices? If, indeed, customer satisfaction is relevant to market valuation, according to efficient markets theory, new information about customer satisfaction fluctuations should be instantly reflected in stock prices. That is, all else being equal, investors should reward firms with increased market capitalization as news about the improved value of firms' customer relationships becomes available. Conversely, news about deteriorating customer relationships should have the opposite effect.

Such reasoning is in concert with financial valuation theory in that customer relationships can be expressed in terms of net present value (Blattberg and Deighton 1996; Hogan et al. 2002; Rust and Keiningham 1994; Venkatesan and Kumar 2004). In competitive markets, the value of satisfied customers should be worth more than the value of dissatisfied customers.

However, the picture is more complicated. There are reasons for suggesting that investors may not react in a predictable fashion, and there are circumstances in which stock prices drop as a result of news about increasing customer satisfaction. Although such a reaction might be at odds with how capitalistic markets should function, consumer market imperfections (which may be unknown) make it difficult to anticipate capital market reactions to customer satisfaction news. First, investors may react negatively to news about rising customer satisfaction if they believe that the firm is giving away too much surplus to the buyers. The aggregate surplus that consumers receive from goods and services is a measure of consumer welfare, but the difference between the maximum price a buyer is willing to pay and the actual price is also something that investors may want to capture. Investors would be more likely to consider an increase in consumer surplus in a negative light if the buyer's switching costs were significant, if there was a high degree of product differentiation, or if there was some level of monopoly power.

Second, investors may also take an unkind view of improvements in customer satisfaction for firms that already have substantial leads over their competition. If customers of competing firms were much less satisfied, the marginal return of improving customer satisfaction further might be called into question. Certainly, it would be difficult to ignore economic laws of diminishing returns for the investment in any economic asset, regardless of its status on the balance sheet.

Third, and related to the preceding two points, the marginal cost of improving customer satisfaction may be too high. This is particularly relevant for services, especially those that are labor intensive. For example, it has been shown (Anderson, Fornell, and Rust 1997) that productivity and customer satisfaction are not always compatible in the service sector. Improved customer satisfaction may come at the expense of productivity, and vice versa. Such a trade-off between productivity and customer satisfaction may also contribute to unfavorable market reactions to news about customer satisfaction improvements.

Fourth, there is the possibility of a "reverse causality effect." Customer defection can have a positive effect on average customer satisfaction simply because the departing customers were the most dissatisfied. Customers who remain are less likely to be as discontented. Instead of benefiting from improved satisfaction, the company simply retains a smaller group of (somewhat more satisfied) customers, but often with reduced sales and profits as a result. Note that this effect is not the result of consumer market imperfections. Indeed, it is more likely to occur in fluid markets with low switching barriers and no monopoly power.

Fifth, issues of timing and expectations complicate the matter further. The ACSI measures each company once a year. Although the measurement is cumulative in the sense that it aims to capture all relevant customer experience (with the product or service) to date, it is possible for changes in customer satisfaction to occur and to have an effect on buyer behavior (with implications on revenues and profits) in periods between measurements. It is also difficult to ascertain market expectations about customer satisfaction. A firm's improved ACSI scores may be viewed as negative if the firm was expected to show a greater degree of improvement, and vice versa. Sources of such expectations may be difficult to identify; they may include corporate announcements about additional investments in consumer service, the hiring of more frontline personnel, investments in information technology and customer relationship management systems (Krishnan et al. 1999; Mithas, Krishnan, and Fornell 2005a, b; Prahalad, Krishnan, and Mithas 2002; Rai, Patnayakuni, and Patnayakuni 2006; Sambamurthy, Bharadwaj, and Grover 2003; Srinivasan and Moorman 2005), and a host of things that may affect customer satisfaction.

For all these reasons, it is not easy to predict stock market reactions to news about customer satisfaction for individual companies. It would not be enough to appeal to capitalistic market theory without consideration of an assortment of market conditions. Even if the relevant characteristics were observed and properly assessed, the effect may be confounded by the shape of the satisfaction response curve, by the firm's location on the curve, or by the reverse causation phenomenon. Thus, even if investors do not react to news about customer satisfaction, share prices may still be unbiased. Only if it is possible to use ACSI information to discover mispriced stocks consistently over time and invest accordingly (and thus beat the market) can the question about possible pricing bias be addressed. Accordingly, we conduct both an event study and portfolio studies. We begin with the event study.

We estimate the rate of return on stock price of firm j on day t with the following market model:

$$R_{jt} = \alpha_j + \beta_j R_{mt} + \varepsilon_{jt}, \tag{2}$$

where R_{jt} is the rate of return on the common stock of the jth firm on day t, R_{mt} is the market rate of return using the Standard & Poor's (S&P) 500 composite index on day t, α_j is an intercept, β_j is a slope parameter that measures the sensitivity of R_{jt} to the market index, and ε_{jt} is a disturbance term with the usual ordinary least squares properties. The S&P 500 is a capitalization weighted index based on a broad cross-section of the market and has been used in several previous event studies (MacKinlay 1997).

We use the market model (Equation 2) to estimate the abnormal return for the common stock of firm j on day t, such that

$$AR_{jt} = R_{jt} - (\alpha_j + \beta_j R_{mt}). \tag{3}$$

On the basis of recommended guidelines and commonly used practice, we specified the length of the estimation period as 255 days, ending 46 days before the event date (Cowan 2003; McWilliams and Siegel 1997).

We averaged daily abnormal returns using the closing price at the day of the release over the sample of N firms to yield cumulative abnormal returns:

$$CAR_{T_1T_2} = \frac{\sum_{j=1}^{N} \sum_{t=T_1}^{T_2} AR_{jt}}{N}, \tag{4}$$

where T_1 and T_2 correspond to the first and last days of the event period.

Because the sampling distribution of abnormal returns tends to be skewed and leptokurtic (Brown and Warner 1985; Corrado 1989; Lyon, Barber, and Tsai 1999; McWilliams and Siegel 1997), we used bootstrapping for all significance tests. Use of bootstrapped test statistics provide a more robust assessment of the statistical significance of the results of the event study and is becoming increasingly common in academic studies (Chatterjee, Richardson, and Zmud 2001).

It is well known that event studies benefit considerably from a precise definition of the event period (Barclay and Litzenberger 1988; Brown and Warner 1985; Dyckman, Philbrick, and Jens 1984). As for the ACSI, there is an exact announcement day. Thus, we specify a one-day event period. A short event period leads to more efficient estimation because it reduces the possibility of other factors affecting the returns. It also increases the power of the statistical tests.

As we indicated previously, we defined the event as the ACSI announcement for a firm as (simultaneously) published in *The Wall Street Journal* and on the ACSI Web site. Because the *The Wall Street Journal* is a daily publication, it seems reasonable to consider the day the ACSI announcement is published in *The Wall Street Journal* as the day when "the event" occurs. The included events range from the second quarter of 1999 to the third quarter of 2002.[2]

Following recommended practice (MacKinlay 1997), we set up controls so that the customer satisfaction news would not be confounded with other news items over the same period. We collected all firm-related news items five calendar days before and after the event date from leading sources: PR Newswire, Dow Jones, and Business Wire. We excluded from the analysis events with announcements regarding mergers and acquisitions, spin-offs, stock splits, chief executive officer and chief financial officer changes, layoffs, restructurings, earnings announcements, and lawsuits. In total, after we eliminated the events with confounded news items during the event windows, the data consisted of 161 events for 89 companies for which stock-trading data were available for the study period. For manufacturing firms, there were 48 events of rising ACSI scores; for service firms, the corresponding number was 32. There were 41 events with decline in ACSI for manufacturing and 40 for services. The results appear in Table 2.

Although the combined samples for manufacturing and services changed in the direction of higher stock prices for firms with an increase in ACSI and lower prices for firms with a decrease in ACSI, these effects are too weak to be distinguished from chance variation. The results for service firms are also contradictory; that is, share prices increased regardless of the direction of the change in ACSI.

These results are fairly consistent with the findings of Ittner and Larcker (1998). They used a longer event window of five days and found no significant effects of ACSI news. For a ten-day window, they reported close to significant (at the 10% level) abnormal returns but only after they deleted outliers (still 97%–98% of the variance in the abnormal returns remained unexplained). Note that Ittner and Larcker were forced to use cross-sectional analysis (because of the limited amount of data available at the time), and thus their results cannot be used to test the effect of ACSI on changes in stock prices (Lambert 1998).[3]

TABLE 2
Abnormal Returns For ACSI News Announcements

Type of News	Firm Type	N	Event Day Abnormal Returns
Increase in ACSI	Manufacturing and services	80	.34
	Manufacturing	48	.01
	Services	32	.85
Decrease in ACSI	Manufacturing and services	81	–.17
	Manufacturing	41	–.54
	Services	40	.20

Notes: None of the abnormal returns are significant at the 5% level based on nonparametric bootstrapped tests (two tailed)

---

[2] Although ACSI has measured customer satisfaction since 1994, before the second quarter of 1999, the results were published once a year in *Fortune* magazine, making it difficult to pinpoint the event date because readers received the magazine on different dates. For example, Ittner and Larcker's (1998) study used a much longer event window (five and ten days) in their analysis of the effect of the *Fortune* publication

[3] Although it seems unlikely that the ACSI information would trickle out slowly or that investors would simply wait before acting on the information, we also calculated cumulative abnormal returns for a 5-day window following the release date. Again, the null hypothesis of zero abnormal returns cannot be rejected. A final test checks for the possibility of prior leakage of information. The ACSI results were routinely provided under embargo to the public relations and market research units of corporate subscribers

Portfolio Studies

The finding that news about customer satisfaction does not move stock prices, even though customer satisfaction is statistically significant with respect to market valuation, is not necessarily contradictory to efficient markets theory. It would still be possible that stock prices fully and quickly reflect all available information. It is also possible that firms with highly satisfied customers generate higher profits. Under efficient markets theory, however, it would not be possible to develop trading strategies based on ACSI information that consistently earn excess returns.

As a way to test this, we consider two stock portfolios based on ACSI data. The first is a hypothetical paper portfolio with simple trading rules. The second is a real-world portfolio. Both tests have pros and cons. However, their combination leads to more powerful analyses. The major advantage of the paper portfolio is that the trading rules are clear and explicit. In addition to there being no theoretical basis for simulating various permutations of the trading strategy, its major disadvantage is that it is hypothetical and back tested with perfect hindsight.

For the real-world portfolio, these advantages and disadvantages are reversed. Actual trading involves many details regarding timing, size of investments, number of holdings, rebalancing, and so forth, that are not always precisely determined by static and explicit decision rules. In many ways, the situation is similar to case studies of organizations for which theoretical specificity must give way to realism. Yet this is also the benefit of a real-world portfolio: It presents authentic investments and actual results.

Most important, however, paper portfolio and the actual portfolio share a critical aspect: Both rely solely on ACSI data. No other type of data or additional information was considered in any of the investment decisions. Although details may vary between the portfolios, and assuming an absence of contemporaneous underlying share price determinants that covary with ACSI (we discuss this further subsequently), it would be difficult to attribute the results to sources other than customer satisfaction.

Because the market does not immediately respond to changes in ACSI, it seems reasonable to base the strategy on both levels (because they do not appear to be fully impounded in the market price) and changes (because they do not appear at all to be impounded in the market price). To have a diversified portfolio of reasonable size, we selected firms in the top 20% of ACSI (relative to their competition). In line with Jones and Sasser (1995) and Ittner and Larcker (1998), who find that low levels of customer satisfaction appear to have less of an impact, we also made the selection conditional on the requirement of being above the ACSI national average.

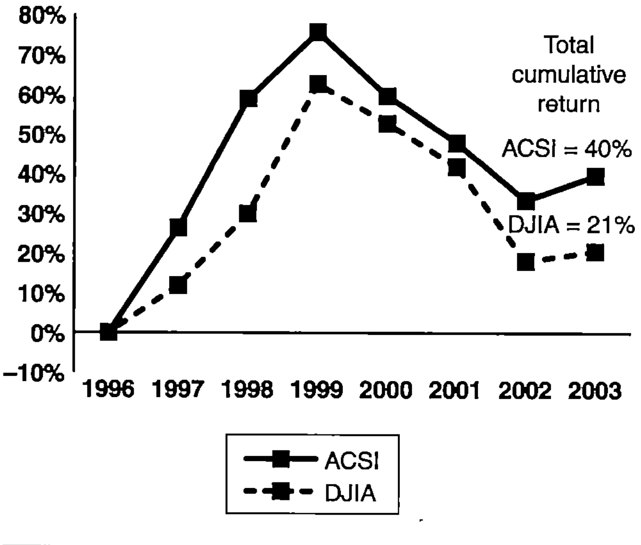
Because ACSI scores for each individual firm are updated once a year, we purchased shares (using the closing price) the day the ACSI results were announced and held them for a year or longer. If we picked a stock the first year, we examined its inclusion in the portfolio again the following year. If it met the criteria of being in the top 20% and above the national (average) ACSI level, we held it for another year and then subjected it to the same test. If it failed the criteria, we sold it. We applied the same principle for all stocks. If we did not pick a stock the first year, we reexamined it for inclusion the next year, and so on. This trading strategy led to a portfolio of 20 companies in 1997 and 20 in 1998, and there were 26 companies at the conclusion of the test on May 21, 2003.

Between February 18, 1997, and May 21, 2003, a period when the stock market had both ups and downs, the portfolio generated a cumulative return of 40% (with dividends and transaction costs excluded but after stock splits, if any, were adjusted). It outperformed the Dow Jones Industrial Average (DJIA) by 93%, the S&P 500 by 201%, and NASDAQ by 335%. Figures 1–3 show the cumulative returns over time.

The results suggest that customer satisfaction pays off in up-markets and down-markets. When the stock market grew, the stock prices of many firms with highly satisfied customers grew even more. The only exception occurred at the peak of the stock market bubble in 1999, when NASDAQ and the S&P 500 generated short-lived, but higher, returns. When the stock market dropped in value, the stock prices of firms with highly satisfied customers seemed to have benefited from some degree of insulation.

Although the trading rules barred all information beyond ACSI, could there be competing explanations for the findings? We examine the possibilities. First, and regardless of whether efficient markets theory holds, it is difficult to beat the market. Most professional stock pickers do not, and the chance probability of consistent excess returns is minute. Second, there are theoretical underpinnings behind these returns that are bolstered by empirical

FIGURE 1
Cumulative Returns: ACSI Top 20% Versus DJIA (1997–2003)[a]

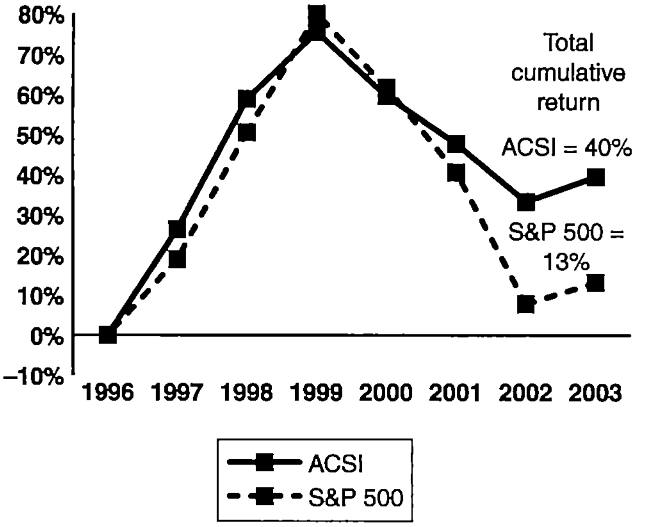


[a]February 18, 1997, through May 21, 2003.

---

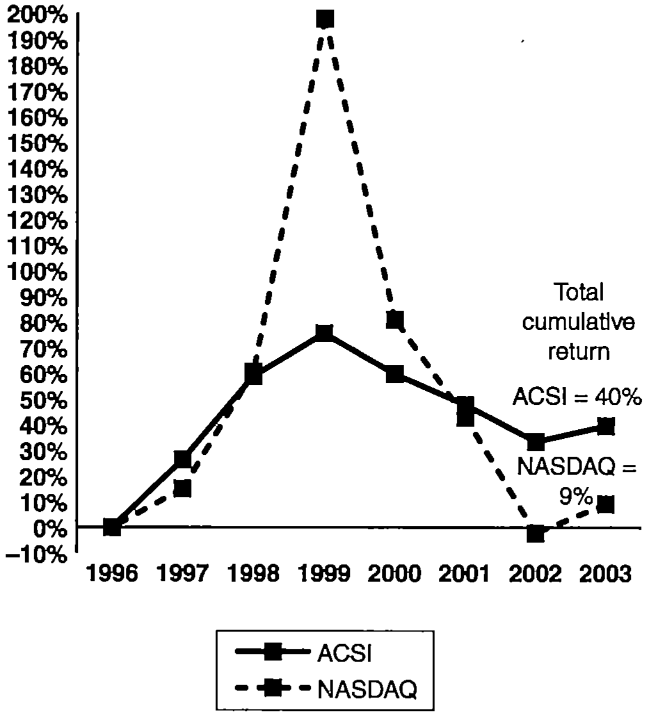
and to *The Wall Street Journal* about two weeks before the release. That corresponds to 10 additional trading days before the event. Again, the results are not significant. The results of 5-day and 15-day analyses appear in the Appendix.

FIGURE 2
Cumulative Returns: ACSI Top 20% Versus S&P 500 (1997–2003)[a]



[a]February 18, 1997, through May 21, 2003.

FIGURE 3
Cumulative Returns: ACSI Top 20% Versus NASDAQ (1997–2003)[a]



[a]February 18, 1997, through May 21, 2003.

findings in the marketing literature. Third, the abnormal returns are not due to compensation for risk. The beta risk associated with the portfolio is .78 and thus is substantially less than market.[4] Fourth, there is no known ACSI covariate that accounts for the results. Firms are included in ACSI on the basis of size (Fornell et al. 1996), and the ACSI aggregate should approximately mirror large cap indexes (e.g., DJIA). There is no reason to suspect that the top 20% of ACSI firms would have the same performance as the bottom 80%. To verify this empirically, we collected stock price data for all ACSI firms. The cumulative return for the bottom 80% portfolio was 20.4%, which is virtually identical to that of the DJIA (21%), thus ruling out selection criteria as a covariate of the stock performance. Because book-to-market ratios are sometimes mentioned as possible predictors of long-term abnormal stock returns (Fama and French 1992), we compared these ratios as well. For the top 20%, the book-to-market ratio was .41 for the full time period. The corresponding ratio for the bottom 80% was .42. Finally, we examined the "size effect." It has been shown that smaller firms' stocks tend to produce higher returns (Banz 1981; Bernard and Thomas 1989; Reinganum 1981). Are the top 20% of ACSI firms smaller than the bottom 80%? Consistent with the results in Table 1, the opposite is true. The average revenue for the top 20% of firms is $37.8 billion, compared with $26.4 billion for the bottom 80%.

Overall, these results are noteworthy because they suggest a trading strategy with high returns and low risk, based on theory and findings not in finance or economics but rather in the marketing literature. Because of the magnitude of the abnormal returns, the findings are perhaps even more remarkable considering that most brokerage firms had negative abnormal returns on their stock recommendations between 1997 and 2001 (McGeehan 2001). Nonetheless, a word of caution is warranted. The results refer to a paper portfolio that is back tested. As with all such testing, there are obvious limitations. In hindsight, it may not be that difficult to find a successful trading strategy. In defense of the results, however, we note that the trading strategy used was extraordinarily simple and straightforward.

However, not all individual trades were good ones. Gateway computers is a case in point. Stock was purchased on August 21, 2000, because the company had an ACSI score of 78, which was in the top 20% of its industry and above the national average. Stock was sold on August 20, 2001, because the company dropped out of the top 20%. Over that period of time, the company's stock price fell by 84% because consumer demand for personal computers declined sharply. As we discussed previously, there may be many reasons that the trading strategy may not work for each individual company all the time. In addition, there is no reason to believe that high customer satisfaction will provide full stock market insulation if primary consumer demand falters. It may dampen the fall to some degree, but industry demand is affected by many other factors. However, from a diversified portfolio, negative industry disturbances should be counterbalanced by positive ones. To some extent, the large losses made on Gateway in one period were compensated by a 47% gain in another period (1999–2000), as well as Dell Computer's return of 50% in 1998–1999 and the very large rise of the Wal-Mart (96% in 1998–1999) and Southwest Airlines (96% in 1998–1999) stocks.

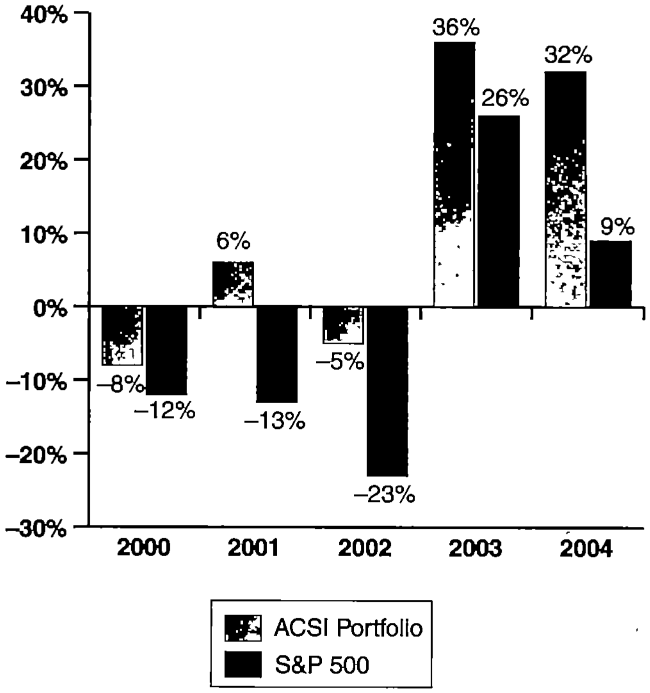
---

[4] The beta values are relative to the S&P 500 over the entire life of the portfolio. On a year-by-year basis, the betas are .72 (1997), .75 (1998), .74 (1999), .75 (2000), .76 (2001), .81 (2002), and .93 (2003).

Our final analysis goes from hypothetical paper trading to an actual stock portfolio of ACSI companies. Thus, it avoids the usual limitations of back testing and provides additional evidence about the strength of the relationship between customer satisfaction and stock price. As before, we considered no other data but ACSI data in the investment decisions. We took both long and short positions and based trading decisions on ACSI levels, changes, and satisfaction elasticity of demand.[5] The portfolio took long positions in firms with high (above competition) and increasing (by two points or more) ACSI scores and also when the ACSI–customer retention (elasticity) estimate was above average. It took short positions when scores were low (lower than any competitor) and deteriorating (by two points or more). The short positions were generally small and ranged over time from 0% to 20% of the portfolio. We held these positions for a one-year period or more. We balanced the portfolio weighting with approximately the same dollar amount for each company at the starting point but also with consideration given to industry representation.[6] We periodically rebalanced to dollar equalize the positions, and we executed trades throughout the year without attention to the date of the ACSI release.[7] We reinvested dividends and included transaction costs. Trading began in April 2000, at a time just before the burst of the stock bubble from which the market, as of the end of 2004, has not yet recovered.

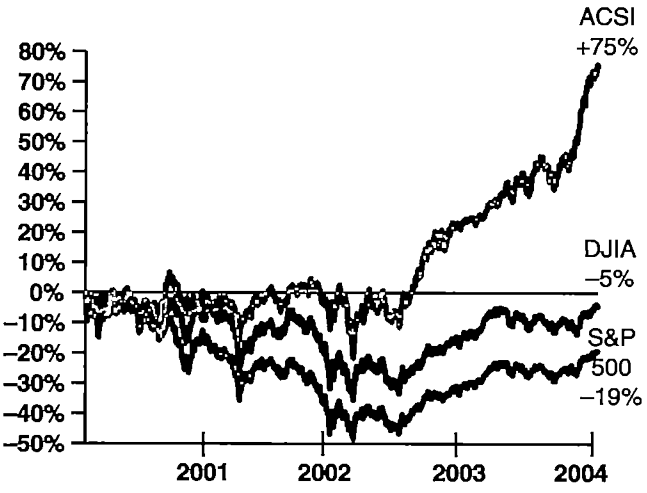
The portfolio generated impressive positive returns. As we show in Figure 4, it outperformed the S&P 500 each and every year during the five years since its inception, often by a considerable margin. The smallest gain relative to market was in 2000 when the S&P dropped by 12% and the ACSI portfolio dropped by 8%.[8] In the following year, the portfolio gained 6%, whereas the S&P continued to fall by 13%. With the exception of 2002, when the S&P fell by 23% and the ACSI portfolio fell by 5%, there have been large gains: 36% in 2003 (versus 26% for the S&P) and 32% in 2004 (versus 9% for the S&P). On a cumulative basis (Figure 5),

FIGURE 4
Yearly Performance of ACSI Portfolio and S&P 500 (2000–2004)[a]



[a]April 11, 2000, through December 31, 2004.

FIGURE 5
ACSI Portfolio Versus DJIA and S&P 500 (2000–2004)[a]



[a]April 11, 2000, through December 31, 2004.

---

[5] For discussions on this type of elasticity, see Fornell (1992) and Anderson, Fornell, and Mazvancheryl (2004).

[6] The ACSI includes many utilities. To compensate, each utility in the portfolio received only half of the dollar amount invested in nonutility companies.

[7] Accordingly, some trades were executed subsequent to the ACSI release date (approximately 70%), and some were executed before that date. Because there was no announcement event effect, the positive returns were not helped by this timing. On the contrary, they were somewhat handicapped. The cumulative return for the portfolio for all the event days (using the closing price of the day of the ACSI release), through February 2003, was –1.6%. The DJIA and S&P 500 did somewhat better at +.5% and +.6%, respectively All trades subsequent to February 2003 were executed after the ACSI release.

[8] The 2000 results do not represent the full year. They reflect the starting date of the portfolio (April 11, 2000).

the ACSI portfolio gained 75%, compared with a loss of 19% for the S&P 500. The DJIA did slightly better at –4%.[9]

As with the previous portfolio study, these returns are not associated with higher risk. The betas for the real-world portfolio are equivalent to the paper portfolio. The mean beta risk for all years was .77 relative to the S&P and .76 relative to the DJIA. On an annual basis, the beta risk of the portfolio was consistently below market.[10]

Similar to the hypothetical paper version, share prices for firms that did well on ACSI rose faster than the market and dropped at a slower rate. However, short selling produced weaker returns (less than 3% of the total), implying asymmetric impacts from high versus low customer satisfaction. However, it would be premature to conclude on the basis of this finding alone that the market rewards high customer satisfaction more than it punishes low satisfaction. Additional studies would be necessary to draw such a conclusion.

Discussion

By any standard, the results of the studies we report herein are extraordinary. Not only do they show that investments based on customer satisfaction produce sizable excess returns, but they also upset the basic financial principle that assets producing high returns carry high risk. The price of any financial asset is determined by the current value of the future cash payments it generates, discounted to compensate for risk and cost of capital. The economic value of satisfied customers seems to be systematically undervalued, even though these customers generate substantial net cash flows with low volatility. Firms that do better than their competition in terms of satisfying customers (as measured by ACSI) generate superior returns at lower systematic risk. Similar security mispricings, but on a much lower scale, have been found in related situations with respect to quality awards announcements (Hendricks and Singhal 2001) and marketing investments (Penman and Zhang 2001), but most discoveries of security mispricings point to “overpricing” (Shiller 2000). Under the assumption of efficient markets theory, if information about customer satisfaction is relevant to a company’s future economic performance, it should be factored into the company’s stock price such that it would not be possible to earn economic profits by trading on this information. Yet we show that this is indeed possible. However, this does not imply a wholesale rejection of efficient markets theory. There have been many documented anomalies over the years (e.g., the January effect, the small-firm effect, the Monday effect, the post-earnings-announcement drift). Most of them have since disappeared. Whether that will be the case with the customer satisfaction effect is yet to be determined, but it is less likely because, with the exception of the public release of ACSI and contrary to the January effect and others, the information is not costless and requires more sophisticated measurement technology than firms typically use today (Ittner and Larcker 2003). Because investments based on customer satisfaction may not be publicly known, it would also be much more difficult to verify a possible demise. According to Gupta, Lehmann, and Stuart (2004), financial analysts have yet to give more than scant attention to off-balance-sheet assets, even though these assets may be key determinants of a firm’s market value. By using a discounted cash flow analysis for estimating the value of customer relationships, Gupta, Lehmann, and Stuart find that some companies were potentially mispriced, whereas others were not.

Taken as a whole, the conclusion from these analyses is that though firms with highly satisfied customers usually generate positive abnormal returns, news about changes in customer satisfaction does not have an effect on stock prices. In other words, the primary thesis behind capitalist markets and marketing theory is borne out, but not without reservations about efficient (equity) markets theory. Specifically, there seem to be imperfections with respect to the time it takes for stock markets to reward firms that do well by their customers and to punish firms that do not. In the wake of accounting scandals, the bursting of the stock market bubble, and the continued weakening of the relationship between balance sheet assets and future income, it would be in the interest of securities research to pay closer attention to customer satisfaction and the strength of customer relationships. For marketing managers, it is clear that the cost of managing customer relationships and the cash flows they produce is fundamental to value creation.

The principal short-term impact of the findings may be to boost security analysts’ demand for customer satisfaction information. If history is a guide, however, it will be difficult for analysts to separate the “good” from the “bad” and the relevant from the irrelevant when it comes to this type of information. This will be a serious challenge because recent experience with valuing intangible assets did not turn out well. Just before the stock market bubble burst, Wall Street seemed enamored with customer metrics, most of which had dubious connections to economic value creation and large losses in shareholder wealth as a result. Thus, a more prudent approach might be expected this time, but appropriate prudence does not suggest eschewing customer metrics altogether, which more or less seems to be the case for securities research today.[11] Rather, the quality of measurement and economic relevance should dictate how these metrics should best be used. For example, although customer satisfaction is strongly related to economic value creation and the (collective) customer may have information relevant to the financial prospects of the firm before investors do, this cannot be leveraged if measurement is flawed and unable to extract the information. Ittner and Larcker (2003) find that most companies' measurement methodologies for customer satisfaction are "mindless," misleading, and too primitive to be useful. This state of affairs is unlikely to change unless shareholders, corporate boards, and investors put more pressure on companies to account for intangible assets more effectively. If accounting were to include information about the health of the firm's customer relationships, investors would have a better understanding of the link between the firm's assets and its capacity to generate shareholder wealth, and managers would be compensated accordingly. Corporate governance would benefit as well.

---

[9] Because the portfolio includes reinvested dividends and transaction costs, the comparison with market indexes is not exactly “apples to apples,” but the gains relative to market are much greater than the net effects from dividends and transaction costs. For example, with dividends and transaction costs included, the average diversified stock portfolio rose only 1% according to Lipper Research over the same time period.

[10] The year-to-year ACSI portfolio betas relative to the DJIA are .87 (2000), .67 (2001), .74 (2002), .82 (2002), and .84 (2004). The year-to-year ACSI portfolio betas relative to the S&P 500 are .83 (2000), .73 (2001), .72 (2002), .82 (2003), and .87 (2004).

[11] This is not to suggest that all securities analysts completely ignore customer satisfaction. For example, Merrill Lynch and Morgan Stanley have written reports on it, and in general, analysts attribute the surge in Apple’s stock price in late 2004 to an increase in customer satisfaction.

If investors were to use capital more expediently to encourage firms to strive for higher levels of customer satisfaction and better quality of measurement, there would be many beneficiaries. Consumers would be better off. As consumer and equity markets become more synchronized, their joint influence may contribute to less misallocation of capital, quicker deflation of stock market bubbles (see Figures 1–3), fewer cases of security mispricings, and a better functioning of markets in general.

Although investors fail to realize the relevance of ACSI news and equity markets exhibit imperfections in this regard, the results of this study affirm the workings of free capitalistic markets with respect to the relationship between consumer utility and the flow of investment capital. Although it takes a while for equity markets and consumer markets to exercise their joint power, the finding that customer satisfaction and stock prices eventually move together suggests that most markets get it right in the long run. They apparently offer meaningful consumer choice, and short-term fluctuations and other exceptions notwithstanding, sellers may take comfort that they will (eventually) be rewarded for treating customers well and that they risk punishment for treating customers poorly.

APPENDIX
Five and Fifteen Days of Cumulative Abnormal Returns (CAR) for ACSI News Announcements

Type of News	Firm Type	N	CAR 5-Day Window	CAR 15-Day Window
Increase in ACSI	Manufacturing and services	80	1.85	–.95
	Manufacturing	48	.98	1.23
	Services	32	3.15	–4.22
Decrease in ACSI	Manufacturing and services	81	–.22	–1.65
	Manufacturing	41	–.78	–2.72
	Services	40	.36	–.55

Notes: None of the cumulative abnormal returns are significant at the 5% level based on nonparametric bootstrapped tests (two tailed).

REFERENCES

Aaker, David A. and Robert Jacobson (1994), "The Financial Information Content of Perceived Quality," *Journal of Marketing Research*, 31 (May), 191–201.

Ambler, Tim, C.B. Bhattacharya, Julie Edell, Kevin Lane Keller, Katherine N. Lemon, and Vikas Mittal (2002), "Relating Brand and Customer Perspectives on Marketing Management," *Journal of Service Research*, 5 (1), 13–25.

Anderson, Eugene W. (1996), "Customer Satisfaction and Price Tolerance," *Marketing Letters*, 7 (3), 19–30.

———, Claes Fornell, and Donald R. Lehmann (1994), "Customer Satisfaction, Market Share, and Profitability: Findings from Sweden," *Journal of Marketing*, 58 (July), 53–66.

———, ———, and Sanal K. Mazvancheryl (2004), "Customer Satisfaction and Shareholder Value," *Journal of Marketing*, 68 (October), 172–85.

———, ———, and Roland T. Rust (1997), "Customer Satisfaction, Productivity, and Profitability: Differences Between Goods and Services," *Marketing Science*, 16 (2), 129–45.

——— and Mary W. Sullivan (1993), "The Antecedents and Consequences of Customer Satisfaction for Firms," *Marketing Science*, 12 (2), 125–43.

Banz, R. (1981), "The Relationship Between Return and Market Value of Common Stocks," *Journal of Financial Economics*, 9 (March), 3–18.

Barclay, M.J. and R.H. Litzenberger (1988), "Announcement Effects of New Equity Issues and Use of Intraday Price Data," *Journal of Financial Economics*, 21 (1), 71–99.

Barth, M.E. and M.E. McNichols (1994), "Estimation and Market Valuation of Environmental Liabilities Relating to Superfund Sites," *Journal of Accounting Research*, 32 (Special Issue), 177–209.

Bell, David, John Deighton, Werner J. Reinartz, Ronald T. Rust, and Gordon Swartz (2002), "Seven Barriers to Customer Equity Management," *Journal of Service Research*, 5 (1), 77–85.

Berger, Paul D., Ruth N. Bolton, Douglas Bowman, Elten Briggs, V. Kumar, A. Parasuraman, and Creed Terry (2002), "Marketing Actions and the Value of Customer Assets: A Framework for Customer Asset Management," *Journal of Service Research*, 5 (1), 39–54.

Bernard, V.L. and J.K. Thomas (1989), "Post-Earnings Announcement Drift: Delayed Price Response or Risk Premium," *Journal of Accounting Research*, 27, 1–36.

Blattberg, Robert C. and John Deighton (1996), "Manage Marketing by the Customer Equity Test," *Harvard Business Review*, 74 (July–August), 136–44.

Bolton, Ruth N. (1998), "A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction," *Marketing Science*, 17 (1), 45–65.

———, P.K. Kannan, and Matthew D. Bramlett (2000), "Implications of Loyalty Program Membership and Service Experiences for Customer Retention and Value," *Journal of the Academy of Marketing Science*, 28 (1), 95–108.

———, Katherine N. Lemon, and Peter C. Verhoef (2004), "The Theoretical Underpinnings of Customer Asset Management: A Framework and Propositions for Future Research," *Journal of the Academy of Marketing Science*, 32 (July), 271–92.

Brown, Stephen J. and Jerold B. Warner (1985), "Using Daily Stock Returns: The Case of Event Studies," *Journal of Financial Economics*, 14 (1), 3–31.

Bursk, Edward C. (1966), "View Your Customers as Investments," *Harvard Business Review*, 44 (May–June), 91–94.

Capon, Noel, John U Farley, and Scott Hoenig (1990), "Determinants of Financial Performance. A Meta-Analysis," *Management Science*, 36 (10), 1143–59.

Chatterjee, Debabroto, Vernon J. Richardson, and Robert W. Zmud (2001), "Examining the Shareholder Wealth Effects of Announcements of Newly Created C.I.O. Positions," *MIS Quarterly*, 25 (1), 43–70.

Corrado, Charles J (1989), "A Nonparametric Test for Abnormal Security-Price Performance in Event Studies," *Journal of Financial Economics*, 23 (2), 385–95.

Cowan, Arnold R. (2003), *Eventus 7.5 User's Guide*. Ames, IA: Cowan Research LC.

Dyckman, T., D. Philbrick, and J.S. Jens (1984), "A Comparison of Event Study Methodologies Using Daily Stock Returns: A Simulation Approach," *Journal of Accounting Research*, 22 (Supplement), 1–30.

Fama, Eugene F. and Kenneth R. French (1992), "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47 (2), 427–65.

Fornell, Claes (1992), "A National Customer Satisfaction Barometer: The Swedish Experience," *Journal of Marketing*, 56 (January), 6–22.

——— (1995), "The Quality of Economic Output: Empirical Generalizations About Its Distribution and Relationship to Market Share," *Marketing Science*, 14 (3), G203–G211.

——— (2001), "The Science of Satisfaction," *Harvard Business Review*, 79 (March), 120–21.

———, Michael D. Johnson, Eugene W. Anderson, Jaesung Cha, and Barbara Everitt Bryant (1996), "The American Customer Satisfaction Index: Nature, Purpose, and Findings," *Journal of Marketing*, 60 (October), 7–18.

——— and Birger Wernerfelt (1987), "Defensive Marketing Strategy by Customer Complaint Management: A Theoretical Analysis," *Journal of Marketing Research*, 24 (November), 337–46.

——— and ——— (1988), "A Model for Customer Complaint Management," *Marketing Science*, 7 (Summer), 271–86.

Garvin, David A. (1988), *Managing Quality: The Strategic and Competitive Edge*. New York: The Free Press.

Gruca, Thomas and Lopo Rego (2005), "Customer Satisfaction, Cash Flow, and Shareholder Value," *Journal of Marketing*, 69 (July), 115–30.

Gupta, Sunil, Donald R. Lehmann, and Jennifer Ames Stuart (2004), "Valuing Customers," *Journal of Marketing Research*, 41 (February), 7–18.

Hendricks, Kevin B. and Vinod R. Singhal (2001), "The Long Run Stock Price Performance of Firms with Effective TQM Programs," *Management Science*, 47 (3), 359–68.

Heskett, J.L., W.E. Sasser Jr., and L.A. Schlesinger (1997), *The Service Profit Chain: How Leading Companies Link Profit and Growth to Loyalty, Satisfaction and Value*. New York: The Free Press.

Hogan, John E., Donald R. Lehmann, Maria Merino, Rajendra K. Srivastava, Jacquelyn S. Thomas, and Peter C. Verhoef (2002), "Linking Customer Assets to Financial Performance," *Journal of Service Research*, 5 (1), 26–38.

———, Katherine N. Lemon, and Ronald T. Rust (2002), "Customer Equity Management: Charting New Directions for the Future of Marketing," *Journal of Service Research*, 5 (1), 4–12.

Ittner, C.D. and D.F. Larcker (1996), "Measuring the Impact of Quality Initiatives on Firm Financial Performance," in *Advances in the Management of Organizational Quality*, Vol. 1, Soumen Ghosh and Donald Fedor, eds. Greenwich, CT: JAI Press.

——— and ——— (1998), "Are Nonfinancial Measures Leading Indicators of Financial Performance? An Analysis of Customer Satisfaction," *Journal of Accounting Research*, 36 (Supplement), 1–35.

——— and ——— (2003), "Coming Up Short on Non-Financial Performance Measurement," *Harvard Business Review*, (November), 88–95.

Jackson, Barbara Bund (1985), *Winning and Keeping Industrial Customers: The Dynamics of Customer Relationships*. Lexington, MA: Lexington Books.

Jones, Thomas O. and W. Earl Sasser (1995), "Why Satisfied Customers Defect," *Harvard Business Review*, 73 (November–December), 88–100.

Keller, Kevin L. (1993), "Conceptualizing, Measuring, and Managing Customer-Based Brand Equity," *Journal of Marketing*, 57 (January), 1–22.

Krishnan, M.S., Venkat Ramaswamy, Mary C. Meyer, and Paul Damien (1999), "Customer Satisfaction for Financial Services: The Role of Products, Services, and Information Technology," *Management Science*, 45 (9), 1194–1209.

Lambert, Richard A. (1998), "Customer Satisfaction and Future Financial Performance: Discussion of 'Are Nonfinancial Measures Leading Indicators of Financial Performance? An Empirical Analysis of Customer Satisfaction,'" *Journal of Accounting Research*, 36 (Supplement), 37–46.

Landsman, W. (1986), "An Empirical Investigation of Pension Fund Property Rights," *Accounting Review*, 61 (4), 662–91.

Lyon, John D., Brad M. Barber, and Chih-Ling Tsai (1999), "Improved Methods for Tests of Long-Run Abnormal Stock Returns," *Journal of Finance*, 54 (1), 165–201.

MacKinlay, A. Craig (1997), "Event Studies in Economics and Finance," *Journal of Economic Literature*, 35 (1), 13–39.

McGeehan, Patrick (2001), "Market Place: Study Questions Advice from Brokerage Firms," *The New York Times*, (May 29), C4.

McWilliams, Abagail and Donald Siegel (1997), "Event Studies in Management Research: Theoretical and Empirical Issues," *Academy of Management Journal*, 40 (3), 626–57.

Mithas, Sunil, Joni L. Jones, and Will Mitchell (2004), "Determinants of Governance Choice in Business-to-Business Electronic Markets: An Empirical Analysis," working paper, Ross School of Business, University of Michigan.

———, M.S. Krishnan, and Claes Fornell (2005a), "Effect of Information Technology Investments on Customer Satisfaction: Theory and Evidence," working paper, Ross School of Business, University of Michigan.

———, ———, and ——— (2005b), "Why Do Customer Relationship Management Applications Affect Customer Satisfaction?" *Journal of Marketing*, 69 (October), 201–209.

Naumann, Earl and Steven H. Hoisington (2001), *Customer Centered Six Sigma: Linking Customers, Process Improvement, and Financial Results*. Milwaukee: ASQ Quality Press.

Penman, Stephen H. and Xiao-Jun Zhang (2001), "Accounting Conservatism, the Quality of Earnings, and Stock Returns," *The Accounting Review*, 77 (2), 237–64.

Prahalad, C.K., M.S. Krishnan, and Sunil Mithas (2002), "Customer Relationships: The Technology Customer Disconnect," *Optimize*, (December), (accessed March 15, 2005), [available at http://www.optimizemag.com/issue/014/customer.htm].

Rai, A., N. Patnayakuni, and R. Patnayakuni (2006), "Firm Performance Impacts of Digitally-Enabled Supply Chain Integration Capabilities," *MIS Quarterly*, forthcoming.

Reichheld, Frederick F. and W. Earl Sasser Jr. (1990), "Zero Defections: Quality Comes to Services," *Harvard Business Review*, 68 (September–October), 105–111.

——— and T. Teal (1996), *The Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value*. Boston: Harvard Business School Press.

Reinartz, Werner J. and V. Kumar (2000), "On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing," *Journal of Marketing*, 64 (October), 17–35.

——— and ———. (2002), "The Mismanagement of Customer Loyalty," *Harvard Business Review*, 80 (July), 86–94.

Reinganum, R.M. (1981), "Misspecification of the Capital Asset Pricing: Empirical Anomalies Based on Earnings Yields and Market Values," *Journal of Financial Economics*, 12 (March), 89–104.

Rust, Roland T., Tim Ambler, Gregory S. Carpenter, V. Kumar, and Rajendra K. Srivastava (2004), "Measuring Marketing Productivity: Current Knowledge and Future Directions," *Journal of Marketing*, 68 (October), 76–89.

——— and Timothy L. Keiningham (1994), *Returns on Quality: Measuring the Financial Impact of Your Company's Quest for Quality*. Chicago: Probus.

———, Katherine N. Lemon, and Das Narayandas (2005), *Customer Equity Management*. Upper Saddle River, NJ: Pearson.

———, Christine Moorman, and Peter R. Dickson (2002), "Getting Return on Quality: Revenue Expansion, Cost Reduction, or Both?" *Journal of Marketing*, 66 (October), 7–24.

Sambamurthy, V., Anandhi Bharadwaj, and Varun Grover (2003), "Shaping Agility Through Digital Options: Reconceptualizing the Role of Information Technology in Contemporary Firms," *MIS Quarterly*, 27 (2), 237–63.

Shiller, Robert J. (2000), *Irrational Exuberance*. Princeton, NJ: Princeton University Press.

Srinivasan, Raji and Christine Moorman (2005), "Strategic Firm Commitments and Rewards for Customer Relationship Management in Online Retailing," *Journal of Marketing*, 69 (October), 193–200.

Srivastava, Rajendra K., Tasadduq A. Shervani, and Liam Fahey (1998), "Market-Based Assets and Shareholder Value: A Framework for Analysis," *Journal of Marketing*, 62 (January), 2–18.

Venkatesan, R. and V. Kumar (2004), "A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy," *Journal of Marketing*, 68 (October), 106–125.

Wayland, Robert E. and Paul E. Cole (1997), *Customer Connections: New Strategies for Growth*: Boston: Harvard Business School Press

Sungwook Min, Manohar U. Kalwani, & William T. Robinson

Market Pioneer and Early Follower Survival Risks: A Contingency Analysis of Really New Versus Incrementally New Product-Markets

Does the first entrant in a new market have a difficult time surviving or do first-mover advantages provide protection from outright failure? This empirical study of 264 new industrial product-markets yields insights into this controversial research topic. The key data analysis insights arise through a comparison of survival risks in markets that were started with a really new product and in those that were started with an incremental innovation. When a pioneer starts a new market with a really new product, it can be a major challenge just to survive. In contrast, in markets started by an incremental innovation, market pioneer survival risks are much lower. Notably, early followers have the same survival risk across both types of markets. Overall, these results indicate that in markets started by a really new product, the first to market is often the first to fail. In contrast, in markets started by an incremental innovation, it appears that first-mover advantages protect the pioneer from outright failure.

Most literature reviews on survivors conclude that market pioneers tend to have a greater market share than later entrants (see, e.g., Kalyanaram, Robinson, and Urban 1995; Szymanski, Troy, and Bharadwaj 1995; VanderWerf and Mahon 1997). However, the typical relationship between order of market entry and survival is controversial. It is important to try to resolve this controversy because market pioneers' survival problems can easily offset their market share advantages, and firms may be discouraged from investing in the costly and risky attempt to pioneer new markets.

Some research concludes that market pioneers experience a greater survival risk than later entrants. Olleros (1986, p. 8) cites examples of unsuccessful pioneers of new technologies and observes, "[A]gain and again, we see industries emerge over the dead bodies of their early pioneers." Lambkin and Day (1989) and Tellis and Golder (1996) predict higher failure rates for market pioneers, and Lilien and Yoon (1990) and Golder and Tellis (1993) provide empirical evidence in support of this viewpoint. For example, Golder and Tellis report a lifetime market pioneer survival rate of only 53%.

Other research indicates that first-mover advantages help protect the market pioneer from outright failure. These studies include Agarwal's (1997) product life-cycle analysis, Shepherd's (1999) venture capitalist survey, and Robinson and Min's (2002) analysis of industrial goods industries. For example, Robinson and Min report a ten-year survival rate of 66% for market pioneers versus 48% for early followers.

Because of these divergent results, it seems likely that under some circumstances, market pioneers have higher survival rates. Under other circumstances, however, they should have lower survival rates. A contingency analysis may yield important insights into the complex relationship between order of market entry and survival. Srinivasan, Lilien, and Rangaswamy (2004) also examine this complex relationship. They report that network externalities have a negative impact on market pioneer survival.

Although network externalities are important, our contingency analysis is based on product innovation newness. Product innovation newness is an important aspect of new product development that has received considerable research attention (for a literature review on innovation types, see Garcia and Calantone 2002). Our contingency analysis compares survival rates for markets that were started with really new products with those that were started with incremental innovations. This is part of a continuing effort to shed light on different challenges that marketing managers face in really new versus incrementally new product-markets (e.g., Schmidt and Calantone 1998; Song and Montoya-Weiss 1998; Urban, Weinberg, and Hauser 1996).

When a pioneer starts a new market with a really new product, the costs and risks are unusually high, which makes survival more difficult. Because early followers can learn from the pioneer's mistakes, they are not as vulnerable to exit. In contrast, an incrementally new product typically has lower costs and risks, which makes it easier for the market pioneer to survive. In addition, a temporary monopoly plus first-mover advantages should yield higher survival rates for the pioneer than for early followers.

Sungwook Min is Assistant Professor, College of Business Administration, California State University, Long Beach (e-mail: smin2@csulb.edu). Manohar U. Kalwani is OneAmerica Professor of Management (e-mail: kalwani@purdue.edu), and William T. Robinson is Associate Professor of Management (e-mail: wtrobins@purdue.edu), Krannert Graduate School of Management, Purdue University.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

We test our contingency analysis with data from two sources. We used the *Thomas Register of American Manufacturers* to identify 264 new product-markets for industrial goods and to track market pioneer (first entrant) and early follower survival over time. We supplement this with an expert survey that estimates product innovation newness for each of the 264 new markets. The expert survey classifies 66 markets as starting with a really new product and 198 as starting with an incremental innovation.

In the data analysis, consistent results surface between the descriptive statistics and a hazard rate model that accounts for both observed and unobserved heterogeneity. The results reveal that market pioneers that launched a really new product had much lower survival rates than pioneers that launched an incremental innovation. For example, the 12-year pioneer survival rate for really new products is only 23% versus 61% for incremental innovations.

A surprising result is that early followers do not have meaningful survival risk differences in really new versus incrementally new product-markets. For example, the 12-year survival rates in really new versus incrementally new product-markets are essentially equal at 38% and 39%, respectively. This indicates that in really new product-markets, early followers that learn from the market pioneer's experience can materially reduce their survival risk.

The research controversy we described can be resolved by examining survival risk differences within really new and incrementally new product-markets. In incremental innovation product-markets, market pioneers have a consistently higher survival rate than early followers. An opposite result occurs in really new product-markets, in which market pioneers have a lower survival rate. Overall, the key research conclusion is that the market pioneer is more likely to be the first to fail when it starts a new market with a really new product but not when it starts the new market with an incremental innovation.

Survival Risk Hypotheses

The hypotheses we set forth in this section address survival risks for market pioneers and early followers. The first two hypotheses compare survival risks across markets that were started by a pioneer's really new or incrementally new product. The third and fourth hypotheses address survival risk within each type of market by comparing market pioneers with early followers.

Really New Versus Incrementally New Product-Markets

To assess product innovativeness, we follow Urban, Weinberg, and Hauser's (1996) dichotomy of really new versus incrementally new products. They state (p. 47), "Really new products shift market structures, represent new technologies, require consumer learning, and induce behavior changes." At the other end of the spectrum, an incremental innovation is designed to satisfy a felt market need and uses an existing technology or refinement of it.

Really new products, such as automobiles and personal computers, use new technologies to elicit new market needs. Agarwal and Bayus (2002) note that really new products are often unreliable and costly when they first appear on the market. Because of high levels of market and technological uncertainties, they face slow acceptance by potential users. A key source of market uncertainty arises from customers' concern about the new product's ability to deliver all its promised benefits. Really new products may also require the use of complementary hardware (e.g., roads for automobiles) or software (e.g., programs for personal computers) technologies, often supplied by an independent third party, whether government or industry (Olleros 1986). Potential customers often delay adoption of a really new product until the complementary technologies become available.

Really new products often entail high technological uncertainty. Potential users tend to be concerned about whether a market will eventually adopt the technical standards embodied in the really new product. In addition, a dominant product design for a really new product can take years to emerge (Christensen, Suarez, and Utterback 1998). When a pioneer is committed to another product design that customers do not want, its survival can be threatened. In summary, the resources required to overcome market and technological uncertainties associated with the adoption of really new products can cause pioneer burnout and eventual exit from the market.

Conversely, incrementally new products involve much lower levels of market and technological uncertainties. Market uncertainty is low because these new products provide incremental benefits relative to existing products. Because the felt need already exists, market research can provide more accurate sales forecasts for incrementally new products than for really new products. Technological uncertainty is limited because incrementally new products tend to be based on a refinement or extension of established technologies. Thus, market pioneers that launch an incrementally new product should face low survival risk.

There is some empirical evidence of high survival risk for really new products. Olleros (1986) cites pneumatic tires, sewing machines, automobiles, typewriters, helicopters, televisions, jet engines, and transistors as notable examples of product-markets in which pioneers did not survive. Such examples tend to be really new products that serve unmet customer needs. The 36 product categories in Golder and Tellis's (1993) work span innovations such as video recorders, microwave ovens, dishwashers, personal computers, and facsimile machines. They claim (p. 64) to have "chosen to identify pioneers in distinctive product categories rather than in narrow categories," which suggests that their sample is more likely to contain really new products.

In contrast, Urban and colleagues (1986), who report a zero pioneer failure rate, analyze a sample that contains "tightly defined" categories, such as liquid laundry detergents, instant freeze-dried coffee, fabric softeners, and antidandruff shampoo. Robinson and Min (2002) report higher five- and ten-year survival rates for market pioneers than for early followers. Their sample mainly includes incremental innovations, such as disposable mops, animal access doors, boat lights, and pickle-making machinery. Thus, the empirical evidence that supports higher pioneer failure rates emphasizes really new product-markets, whereas the empirical evidence that supports lower pioneer failure rates emphasizes incrementally new markets. Thus:

> H_1: Market pioneers have a greater survival risk for really new products than for incrementally new products.

Market and technological uncertainties associated with the adoption of really new products often take years to resolve. Complementary technologies are likely to be limited until product sales reach a critical mass. The cost and risk of adopting a really new product is often very high. Such breakthrough innovations tend to be relatively expensive and deficient in performance. In light of these deficiencies and uncertainties, many potential users adopt a wait-and-see attitude. The full resolution of market and technological uncertainties may take a long time. Thus, early followers should also experience greater resistance in gaining customer acceptance in really new product-markets.

> H_2: Early followers have a greater survival risk for really new products than for incrementally new products.

Order of Market Entry

Is the first to market the first to fail? A comparison of market pioneer with early follower survival rates in really new and incrementally new product-markets addresses this controversial research topic. For really new products, a market pioneer faces numerous first-mover disadvantages. Lieberman and Montgomery (1988) note first-mover disadvantages as (1) free rider effects, (2) market and technological uncertainties, (3) changing technology or customer needs, and (4) incumbent inertia. Because early followers in our sample typically entered within two to three years of the pioneer, they are also vulnerable to incumbent inertia. Next, we highlight the other first-mover disadvantages.

Imitation is typically less expensive than innovation, so early followers have the opportunity to free ride on the pioneer's investments. This should be especially true for really new products for which pioneering can be costly. As mentioned previously, market and technological uncertainties are often high for really new products, which can require huge resources and cause pioneer burnout (Olleros 1986). Furthermore, Kerin, Varadarajan, and Peterson (1992, p. 43) suggest that "the greater the [market] uncertainty level, the lower the likelihood that a first-mover will make sizable investments in capacity to achieve scale-dependent cost advantages."

Finally, changing technology and customer needs can provide early followers with an opportunity to leapfrog the pioneer with a superior product offering. Product superiority can arise from superior technology or from a superior job of satisfying unmet customer needs.

Two industry studies provide empirical evidence for really new product-markets. Across five diagnostic imaging product categories, Mitchell (1991) finds that later entrants have lower survival risks, and in the rigid disk industry, Christensen, Suarez, and Utterback (1998) report that later entrants have lower survival risks. On the basis of this theoretical and empirical evidence, we predict the following:

> H_{3a}: For really new products, market pioneers have a greater survival risk than early followers.

Although the disadvantages for market pioneers are especially strong for really new products, market pioneer advantages can also be especially strong. First, market pioneers benefit from a temporary monopoly, which provides temporary protection from competition. A temporary monopoly also helps pioneers strengthen their first-mover advantages (Brown and Lattin 1994; Huff and Robinson 1994). Because imitation should be more difficult for really new products than for incrementally new products, this tends to lengthen the pioneer's monopoly.

Second, first-mover advantages help protect pioneers from outright failure. Pioneer first-mover advantages span numerous sources. These sources include a stronger brand name (Schmalensee 1982), customer preferences that are shaped to favor pioneers (Carpenter and Nakamoto 1989), and broader product lines (Robinson and Fornell 1985). In really new product-markets, it seems likely that these first-mover advantages will be stronger. For example, a really new product is more unique than an incremental innovation, and this uniqueness should strengthen consumer memory and brand name recall. Unique products should be more effective in shaping customer preferences in favor of the pioneer. Finally, unique products should have more opportunities for product line extensions. When longer pioneer lead times and stronger first-mover advantages outweigh first-mover disadvantages, we predict the following:

> H_{3b}: For really new products, market pioneers have a lower survival risk than early followers.

For incrementally new products, first-mover disadvantages should be modest. First, there should be fewer opportunities to free ride on the pioneer's investments. This is because these investments are typically modest compared with developing a really new product. Second, incremental innovations typically refine and modify existing products, so the felt need for these products often exists, and the technologies are well established. With established customer needs and technologies, market and technological uncertainties are modest. Finally, changes in customer needs and technologies are less likely to arise and to surprise the pioneer.

In addition, market pioneers should still benefit from a temporary monopoly and first-mover advantages. Even an incremental innovation cannot typically be copied overnight, so the temporary monopoly makes it easier to survive. Furthermore, first-mover advantages in terms of brand name strength, customer preferences that are shaped to favor the pioneer, and broader product lines should make it easier to survive.

Empirically, Robinson and Min (2002) and Urban and colleagues (1986) report relatively high market pioneer survival rates. Because both studies have relatively narrow market boundaries, they should be dominated by incremental innovations. Thus:

> H_{4a}: For incremental innovations, market pioneers have a lower survival risk than early followers.

Although the pioneer disadvantages we discussed previously are not as great in incrementally new product-markets, Golder and Tellis (1993) describe three additional disadvantages. First, improper positioning arises when the pioneer misses the ideal point and when repositioning costs are high. Even for incrementally new product-markets, later entrants have a better chance of hitting the ideal point. Second, resource requirements can change over time. An early follower is in a better position to predict the key resource requirements for survival. Third, a market pioneer may not be able or willing to commit the resources necessary to succeed. Again, an early follower is in a better position to understand the resources required for success. Thus, when the pioneer disadvantages outweigh the advantages, we predict the following:

> H_{4b}: For incremental innovations, market pioneers have a greater survival risk than early followers.

Data

The *Thomas Register* data have been used in numerous survival studies (in the economics literature, see, e.g., Agarwal 1997; Agarwal and Gort 1996, 2001; Gort and Klepper 1982; in the marketing literature, see, e.g., Agarwal and Bayus 2002; Robinson and Min 2002). The *Thomas Register* is a national buying guide that provides "a comprehensive detailed guide to the full range of products manufactured in the United States" (Lavin 1992, p. 129). The *Thomas Register* attempts to achieve comprehensive coverage by subscribing to a broad range of industry newsletters, searching for start-up ventures in university incubators, and providing a free listing in each annual issue. For example, the 2003 *Thomas Register* includes approximately 173,000 firms and more than 65,000 product categories.

A key *Thomas Register* data strength is that the market boundaries are buyer driven. To help industrial buyers, a new product category is typically added to a group of closely related categories. A new category does not split an existing category. For example, an industrial buyer searching for valves will find air valves, nuclear valves, frostproof valves, solar collector drain-down valves, and so forth.

The *Thomas Register*'s professional staff determines market boundaries and the relevant competitors in each market. The market boundaries are narrower than those of Golder and Tellis (1993), who include broad categories, such as microwave ovens, cereals, and soft drinks. However, they are typically broader than the boundaries of Urban and colleagues (1986), who examine consumer packaged goods, such as light beer, antacids, and steak sauce.

Because the *Thomas Register* is a national buying guide, it typically excludes firms with exclusively local sales. By excluding local markets, our sample emphasizes regional and national markets. In terms of scale of entry, this operational definition falls between that of Urban and colleagues (1986), who require national distribution, and that of Golder and Tellis (1993), who include local markets.

We identified new markets by comparing annual editions of the *Thomas Register* from 1960 to 1995. We identified a new market when it was listed in one year but not in prior years. The identification of the 264 new industrial markets in our study was the most time-consuming part of the data collection process.

Thomas Register *Data Limitations*

Although it is unlikely that the *Thomas Register* staff set market boundaries that are too broad, what happens when the market boundaries are too narrow? In these markets, our sample misses the market pioneer that has already entered a broader product-market. Thus, our research inadvertently compares an early follower with a later entrant or even a later entrant with other later entrants. In this situation, the true results should be even stronger than those we subsequently report. This is because the importance of order of entry should have diminishing marginal returns. For example, greater differences should arise between the market pioneer and an early follower than between two later entrants.

Another data limitation arises from our definition of "exit." We follow Agarwal's (1997, p. 574) guidelines such that an exit arises when the firm's name and address are both missing from the product category in a given year. This avoids the problem of confusing an exit with a name change, such as a conglomerate merger, or confusing an exit with an address change from office relocation. Even so, when a successful firm is merged or acquired, its name and address are often changed. In this case, it is classified as an exit. Because exit implies failure rather than success, this limitation surfaces in the other *Thomas Register* survival studies we previously cited.

In an attempt to quantify this potential bias, we gathered additional data for the most recent *Thomas Register* exits that arose between 1994 and 2002. There were 178 product exits during this nine-year period (see Table 1). We then identified cases in which a firm exit arose in the same year as the product exit. A firm exit arises when a firm is delisted from the *Thomas Register*. This approach assumes that when a merger or acquisition occurs, a product exit and the firm exit arise in the same year.

In Table 1, only 25% of the product exits were accompanied by a firm exit. Furthermore, there were no meaningful differences associated with either order of entry or product innovation newness. We provide some additional insights for the 45 cases for which a product exit and the firm exit arise in the same year. We base these insights on a Google search, an Internet search for bankruptcy court filings, and calls to the most recent telephone number listed in the *Thomas Register*.

For these 45 cases, 10 were a result of either a merger or an acquisition, 4 were a result of bankruptcies, and 31 had an unknown outcome.[1] For the mergers and acquisitions, it is important to know whether the acquired firm was successful or unsuccessful in the product category that was being tracked. Successful product category performance should not be classified as an exit, whereas unsuccessful performance can be classified as an exit.

Although there are not enough data to determine success at the time of exit, we assume that when product cate-

---

[1] For the 31 cases with unknown outcomes, we were unable to contact the acquired firm by calling the most recent telephone number listed in the *Thomas Register*. Because successful firms should attempt to maintain telephone contact with established customers, this outcome seems more likely to arise for unsuccessful firms.

TABLE 1
Product and Firm Exits, Mergers and Acquisitions, and Bankruptcies Between 1994 and 2002

	Number of Product Exits	Number (Percentage) of Product Exits That Were Accompanied by a Firm Exit	Mergers or Acquisitions		Number of Bankruptcies	Number of Untraceable Cases
			Survivors	Exits		
Really New Products						
Market pioneers	16	4 (25)	0	1	0	3
Early followers	36	8 (22)	1	0	2	5
Incrementally New Products						
Market pioneers	44	12 (27)	1	1	2	8
Early followers	82	21 (26)	0	6	0	15
Total	178	45 (25)	2	8	4	31

gory performance is successful, the acquiring firm should remain active in the product category through 2002 (roughly three years after the acquired firm's exit). In Table 1, the acquiring firm remained active in the product category for only two of the ten mergers and acquisitions. This suggests that even when a *Thomas Register* exit arises from a merger or an acquisition, the acquired firm is more likely to be struggling than successful. Overall, these results indicate that mergers and acquisitions of successful firms do not seriously bias the *Thomas Register* exit measure.

Key Definitions

In each product-market, the market pioneer is defined as the market's first entrant. Because our research compares market pioneers with early followers, our analysis excludes markets with multiple first-year entrants. This is because a unique first entrant (market pioneer) cannot be identified. When we identified a unique first entrant, we tracked the market forward on an annual basis until we identified another entrant. We call this entrant and any other entrants in that year early followers.

We surveyed industry and research experts to estimate the degree to which each new product-market was started with a really new product or an incremental innovation. We provided each expert with Urban, Weinberg, and Hauser's (1996) conceptual distinctions between really new and incrementally new products. In these conceptual distinctions, a really new product creates a new market, relies on new technology, and requires customer learning. An incrementally new product is designed to satisfy a felt market need and uses an existing technology or refinement of it. In the rating scale designed to elicit the newness of a product innovation, 1 is an incremental innovation and 7 is a really new product. Table 2 provides the survey questionnaire for chemicals and allied products.

For each of the 264 new markets, four to nine experts rated the degree of product innovativeness when the new market started. Of the 39 experts, 15 had a doctoral degree, most had industry work experience, and all were pursuing scientific research. Approximately 90% were employed in a research-oriented university that has most of its engineering departments ranked in the top 20 nationally. Each expert evaluated products only in his or her field of expertise.[2]

To the best of our knowledge, the experts' ratings were internally consistent across the product categories. Even so, in the questionnaire pretest, some experts in basic research labs gave consistently low ratings for the degree of product innovativeness. To control for this bias, our final survey included some well-known really new products that served as benchmarks, which we derived from prior research studies. For chemical and allied products, for example, our survey included DDT, saccharin, penicillin, and streptomycin (Gort and Klepper 1982). In other categories, we included the personal computer, the laser jet printer, the Walkman, Web TV (Urban, Weinberg, and Hauser 1996), transistors, electric shavers, and radar detectors (Gort and Klepper 1982).

To adjust for individual response bias, in the first step, we computed each respondent's benchmark score, which is the average score across the well-known really new products. In the second step, we subtracted the respondent's benchmark score from the corresponding raw score. For each expert, we considered a product really new when the raw score was greater than or equal to the benchmark score. We classified a product as really new when the majority of experts agreed that a given product was really new. This procedure yielded 66 really new products and 198 incremental innovations.

To quantify the reliability of the product innovation newness measure, we determined Cronbach's alpha, which estimates the average correlation in interexpert ratings. Table 3 reports the results across two-digit Standard Industrial Classification (SIC) code industries. Cronbach's alphas ranged from .76 to .96, which indicates an acceptable degree of reliability.

---

[2] The surveys were conducted in person. Although it was not part of the formal survey, in most cases, the expert could not identify the market pioneer or the first early follower. This may arise because the average year of market inception was in the 1970s. Thus, it does not appear that early entrant survival knowledge biased the degree of product innovativeness measure.

TABLE 2
Survey Questionnaire for Chemicals and Allied Products

Please evaluate whether each of the following products was a *really new product* or an *incremental new product* at the time of its initial commercial introduction. Note that:

- A *really new product*, or basic innovation, creates a new market, relies on new technology, and requires customer learning.
- An *incremental new product*, or modification, is designed to satisfy a felt market need and uses existing technology or refinement of it.

Product Category	Incremental						Really New	
Carbon monoxide gas	1	2	3	4	5	6	7	Don't know
DDT	1	2	3	4	5	6	7	Don't know
Dielectric gas	1	2	3	4	5	6	7	Don't know
Chrome dyes	1	2	3	4	5	6	7	Don't know
Indium compounds	1	2	3	4	5	6	7	Don't know
Zinc hardeners	1	2	3	4	5	6	7	Don't know
Lanthanides	1	2	3	4	5	6	7	Don't know
Saccharin	1	2	3	4	5	6	7	Don't know
Bromide lead	1	2	3	4	5	6	7	Don't know
Lead silico fluoride	1	2	3	4	5	6	7	Don't know
Streptomycin	1	2	3	4	5	6	7	Don't know
Lead titanate	1	2	3	4	5	6	7	Don't know
Lithium acetate	1	2	3	4	5	6	7	Don't know
Lithium niobate	1	2	3	4	5	6	7	Don't know
Magnesium perchlorate	1	2	3	4	5	6	7	Don't know
Penicillin	1	2	3	4	5	6	7	Don't know
Manganous sulfate	1	2	3	4	5	6	7	Don't know

Notes: This is a part of the questionnaire for chemicals and allied products. The actual survey questionnaire includes 34 chemical and allied products.

TABLE 3
Interexperts Reliability: Cronbach's Alpha

Survey Group Industry (Two-Digit SIC Code)	Number of Markets	Number of Raters	Cronbach's Alpha
Textile mills products (22), apparel and fabrics (23), lumber and wood products (24), paper and allied products (26), printing, publishing, and allied industries (27), and other industrial products	19	4	.95
Chemicals and allied products (28)	25	9	.87
Materials and fabricated metal products (30, 32, 33, 34): Subgroup 1	40	4	.84
Materials and fabricated metal products (30, 32, 33, 34): Subgroup 2	16	4	.96
Machinery and computer equipment (35) and other industrial products: Subgroup 1	61	8	.90
Machinery and computer equipment (35) and other industrial products: Subgroup 2	24	4	.93
Electronic and electrical equipment (36): Subgroup 1	24	7	.88
Electronic and electrical equipment (36): Subgroup 2	12	5	.86
Measuring, analyzing, and controlling equipments (38): Subgroup 1	18	8	.86
Measuring, analyzing, and controlling equipments (38): Subgroup 2	25	5	.76
Average	26.40	5.64	.88

Overall, the benchmark adjustment to the product innovativeness measure appears to reduce individual response bias. This is because it yields a reliable measure and tends to strengthen the hazard rate modeling results we discuss subsequently.

Descriptive Statistics

Across the entire sample, the market pioneer 12-year survival rate is 52%. Golder and Tellis (1993) report a lifetime market pioneer survival rate of 53%. Does this imply that the first to market is the first to fail? In our entire sample, the 12-year early follower survival rate is only 39%. Because market pioneers have significantly higher survival rates, it appears that first-mover advantages help protect pioneers from outright failure. Although this conclusion applies to the entire sample, more accurate insights arise through a comparison of really new product-markets with incremental innovations.

In our sample, the really new product-markets have 66 market pioneers and 146 early followers. The incremental innovation product-markets have 198 market pioneers and 340 early followers. The average year of market inception is 1978 for really new products and 1975 for incremental innovations. Average pioneer lead time, which is the length of the pioneer's monopoly, is only 2.5 years.

Table 4 describes the sample in terms of two-digit SIC codes and new product newness. The really new products include 13 chemical products, such as gelling agents, cellulose propionate, indium compounds, dielectric gas, parylene, and so forth. Really new products in other industries include voice-activated microphones, osmometers, blood pressure monitors, video microscopes, bomb detectors, and computer screen projectors. The incremental product-markets include welding nozzles, poultry chillers, asphalt concrete, boat coolers, bottle sorters, fiberglass washers, boat loaders, and ice-dicing machinery. The Appendix provides a complete list of the product-markets, the year of market inception, and the degree of product innovativeness.

Using 4-, 8-, and 12-year survival rates, in Table 5, we provide initial hypothesis-testing insights. We do not use longer time horizons, because right-censoring problems limit the sample size. Twelve-year survival rates provide the most information, so we highlight them.

For H_1, 12-year survival rates for market pioneers are much lower for really new products than for incremental innovations. At 23% versus 61%, incremental innovation survival rates are almost three times higher than those of really new products. This difference can be explained by the greater market and technological uncertainties that pioneers must overcome when they launch a really new product. This result provides strong support for H_1.

For H_2, it is surprising that there are no meaningful survival rate differences for early followers in really new versus incrementally new product-markets. This result is also surprising because these early followers typically entered the new product-markets only two to three years after the pioneer. This suggests that early followers benefit from the efforts and experience of the pioneer in overcoming market and technological uncertainties.

For H_3, market pioneers in really new product-markets have significantly lower 12-year survival rates than early followers (23% versus 38%). For H_4, an opposite sign arises: Market pioneers have significantly higher 12-year survival rates (61% versus 39%). The higher levels of market and technological uncertainties in really new product-markets, which should decrease market pioneer survival rates, can explain this sign reversal.

Two industry examples help illustrate these results. First, in our expert survey, a voice-activated microphone was classified as a really new product. It was pioneered in 1992 by Sherwood Communications Associated Ltd., which entered the market as a start-up company with approximately $1 million in assets. The company exited the market in 1993, which is the same year the first early follower, Telephonics Corporation, entered. Telephonics was an industry incumbent and had experience manufacturing headsets and microphones. Its asset size was greater than $250 million. Telephonics survived through 2002 and was an active player in the market.

A second example is a welding nozzle, which was classified as an incrementally new product in our expert survey. It was pioneered in 1987 by Maryland Ceramic and Steatite Co., which was an industry incumbent with more than $1 million in assets. In 2002, it was one of 22 market survivors. The first early follower was Lurmark Ltd., a start-up company. It entered the market in 1990 but exited in 1993.

Hazard Rate Analysis

Although descriptive statistics are useful, a multivariate hazard rate analysis provides more accurate hypothesis-

TABLE 4
Sample by Two-Digit SIC Code (n = 750)

	Really New Products	Incrementally New Products	All
Industry (Two-Digit SIC Code)			
Textile mills products (22)	0	3	3
Apparel and fabrics (23)	0	3	3
Lumber and wood products (24)	0	5	5
Paper and allied products (26)	0	4	4
Printing, publishing, and allied industries (27)	0	2	2
Chemicals and allied products (28)	13	12	25
Rubber and plastics products (30)	3	19	22
Stone, clay, glass, and concrete products (32)	0	8	8
Primary metal industries (33)	2	5	7
Fabricated metal products (34)	2	17	19
Machinery and computer equipment (35)	13	62	75
Electronic and electrical equipment (36)	12	24	36
Measuring, analyzing, and controlling equipments (38)	17	26	43
Other industries	4	8	12
Total	66	198	264
Number of Firms			
Market pioneers	66	198	264
Early followers	146	340	486
Total	212	538	750

KP2610

TABLE 5
Descriptive Statistics

	N	4-Year Survival	8-Year Survival	12-Year Survival
Market Pioneers (H_1)				
Really new products	66	76%	53%	23%
Incrementally new products	198	92%	79%	61%
|t-statistics|		3.72**	4.17**	5.62**
Early Followers (H_2)				
Really new products	146	74%	50%	38%
Incrementally new products	340	76%	52%	39%
|t-statistics|		.45	.47	.32
Really New Products (H_3)				
Market pioneers	66	76%	53%	23%
Early followers	146	74%	50%	38%
|t-statistics|		.27	.41	2.17*
Incrementally New Products (H_4)				
Market pioneers	198	92%	79%	61%
Early followers	340	76%	52%	39%
|t-statistics|		5.53**	6.64**	5.05**

*$p < .05$
**$p < .01$.
Notes. We assume that right-censored observations at the 10th or 11th year survive for 12 years.

testing results. A hazard rate is the conditional likelihood of a firm's failure in time period t, given that the firm has survived through time period (t − 1). Two widely recognized benefits of a hazard rate analysis are that it (1) represents the impact of time-varying explanatory variables on a firm's likelihood of failure and (2) addresses the right-censoring problem (Helsen and Schmittlein 1993, p. 397). Right censoring occurs when a firm does not have a completed spell within the data observation window and when all that is known is that its survival duration lasted from the time of its market entry until the end of the observation window. Because our data are annual observations, we use a discrete hazard rate model that incorporates the impact of time-varying covariates (e.g., competitive firm density in a product-market) on a firm's likelihood of failure. We model the impact of time on a firm's likelihood of failure nonparametrically to capture any arbitrary pattern in the firm's baseline hazard rate over time (see, e.g., Han and Hausman 1990; Meyer 1990). In other words, we assume that the baseline hazard rate function is a step function of time that is constant within a year but is allowed to vary from one year to another. An advantage of this approach is that it enables consistent estimation of the effect of explanatory variables for arbitrary underlying baseline hazard functions. Finally, the hazard rate model incorporates the impact of unobserved heterogeneity on firm hazard rates.

In the hazard rate model, let t denote the survival duration of an entrant with the cumulative probability distribution function F(t).[3] The associated hazard rate and survival functions are h(t) and S(t) = 1 − F(t). Now, consider that we observe firm i through a total of t_i time periods (years in our case), where t_i can range from one year to the entire length of the data observation window. That is, assume that firm i has survived through the year $(t_i - 1)$ and that we observe it into the t_ith period $[t_i - 1, t_i)$ to determine whether it failed during the year or was right censored during the period. If firm i exits during the t_ith year, the probability of its failure in the t_ith period is $[S(t_i - 1) - S(t_i)]$. However, if it was censored during the t_ith year, all that can be determined is that the firm survived through the end of period $(t_i - 1)$; thus, the probability of that event is $S(t_i - 1)$, under the assumption that censoring occurs at the beginning of the interval. Therefore, we can write the likelihood function for firm i, which we are able to observe into the t_ith year, as follows:

$$(1) \qquad L(t_i) = [S(t_i - 1) - S(t_i)]^{1 - \delta_i} \times [S(t_i - 1)]^{\delta_i},$$

where the indicator variable δ_i takes the value of one if firm i is right censored during the t_ith year and zero if firm i exits the market during the t_ith year. Multiplying and dividing the right-hand side of Equation 1 by $[S(t_i - 1)]^{\delta_i - 1}$, we obtain

$$(2) \qquad L(t_i) = \left[\frac{S(t_i - 1) - S(t_i)}{S(t_i - 1)}\right]^{1 - \delta_i} \times [S(t_i - 1)]$$

$$= [h(t_i)]^{1 - \delta_i} \times [S(t_i - 1)],$$

where $h(t_i)$ is the hazard rate function representing the chance of firm i's exit in the t_ith period, given that it has survived through the period $(t_i - 1)$. We now explain our derivation of the expression for hazard rate function, $h(t_i)$. We then obtain an expression for the associated survival function $S(t_i - 1)$.

---

[3] Vanhuele and colleagues (1995) use the same hazard rate model to analyze marriage dissolution data. Their article has a more complete hazard rate model description.

Overall, the hazard rate of firm i in time period t_i depends on the baseline hazard rate function and the firm's covariate values in time period t_i, $X_i(t_i)$. Firm i's hazard rate in time period t_i can be expressed as

$$h[t_i; X_i(t_i)] = \lambda_0 \times \varphi(t_i) \times \phi[X_i(t_i)], \tag{3}$$

where λ_0 is the base hazard rate, $\varphi(t_i)$ is a step function that captures time (or age) dependence of the hazard rate, and $\varphi[X_i(t_i)]$ is a time-varying covariate function. As we mentioned, the age dependence function $\varphi(t_i)$ is modeled as a step function that remains constant within a year equal to, for example, C_t in year t but is allowed to vary from year to year. The time dependence component of the hazard rate function, $\varphi(t_i)$, in year t_i is

$$\varphi(t_i) = \exp[C_{t_i} D(t_i)], \tag{4}$$

where the dummy variable, $D(t_i)$, equals one in year t_i and zero in all other years. No dummy variable is used for the first year because it is not possible to identify both λ_0 and C_1. Therefore, λ_0 is the time dependence component of the hazard rate function in the first year. Positive (negative) estimates of the time-varying C_t coefficients for the other years indicate a higher (lower) hazard rate than the first-year hazard rate, λ_0.

A firm's hazard rate also depends on its characteristics (e.g., its order of entry in a product-market) and product-market factors, some of which may vary with time (e.g., competitive firm density). The impact of covariates on firm i's hazard rate in period t_i is

$$\phi[X_i(t_i)] = \exp[\beta X_i(t_i)], \tag{5}$$

where $X_i(t_i)$ is a vector of (possible) time-varying covariates and β is a vector of the covariate coefficients. A positive estimate of a β coefficient implies that an increase in the value of the covariate will enhance the likelihood of firm failure. Specifically, when the jth covariate increases by one, the conditional probability of firm failure changes by $[\exp(\beta_j) - 1] \times 100\%$.

Substituting the terms $\varphi(t_i)$ and $\phi[X_i(t_i)]$ from Equations 4 and 5 in Equation 3, we obtain

$$h[t_i; X_i(t_i)] = \lambda_0 \exp[C_t D(t_i) + \beta X_i(t_i)]. \tag{6}$$

The survivor function associated with the hazard rate in Equation 6 is given by (Lancaster 1990; Yamaguchi 1991)

$$S(t_i) = \exp[-H(t_i)], \tag{7}$$

where

$$H(t_i) = \int_0^{t_i} h(u)du. \tag{8}$$

Although the time-varying covariates (e.g., competitive firm density in a product-market) can change from year to year, we assume that they remain constant within each year. Given this simplifying assumption, Vanhuele and colleagues (1995) show that Equation 8 can by rewritten as

$$H[t_i; B_i(t_i)] = \lambda_0 B_i(t_i), \tag{9}$$

where $B_i(t_i) = \Sigma_{t=1}^{t_i} \exp[\beta X_i(t) + C_t D(t)]$. Substituting Equation 9 into Equation 8 yields

$$S[t_i, B_i(t_i)] = \exp[-\lambda_0 B_i(t_i)]. \tag{10}$$

We can now substitute the survival function from Equation 10 into the expression for the likelihood function in Equation 1 and obtain the log-likelihood function for the N firms in our sample (see Vanhuele et al. 1995):

$$LL = \sum_{i=1}^{N} \Big((1 - \delta_i) \log\{\exp[-\lambda_0 B_i(t_i - 1)] - \exp[-\lambda_0 B_i(t_i)]\} - \delta_i \lambda_0 B_i(t_i - 1)\Big). \tag{11}$$

The final step in our hazard rate analysis is to account for unobserved heterogeneity. A firm's likelihood of failure may depend on some factors that are not measurable and others for which data are not available. Thus, firms may have different mean survival rates even when their observable characteristics are the same. Unobserved heterogeneity in the base hazard rate is modeled by allowing λ_0 to vary across the population according to a time-invariant distribution. Thus, the likelihood function contributions, which we previously derived, are conditional on a specific value of λ_0 (see Equation 11). We obtain the unconditional likelihood function by weighting the conditional likelihood by the relative occurrence of λ_0. This is specified as

$$L(t_i) = \int_0^{\infty} L(t_i | \lambda_0) g(\lambda_0) d\lambda_0, \tag{12}$$

where $g(\lambda_0)$ is called the (unobservable) mixing distribution. Because of the gamma distribution's flexibility, it is often used as the mixing distribution (Heckman and Singer 1984; Lancaster 1979). Vanhuele and colleagues (1995) show that assuming a gamma mixing distribution leads to a closed-form solution for the log-likelihood function:

$$LL = \sum_{i=1}^{N} \log\left((1 + \delta_i) \frac{\alpha^r}{[B_i(t_i - 1) + \alpha]^r} - \frac{\alpha^r}{\left\{B_i(t_i - 1) + (1 - \delta_i)\exp[\beta X_i(t_i) + C_{t_i} D(t_i)] + \alpha\right\}^r}\right), \tag{13}$$

where α and r are the scale and shape parameters of the gamma mixing distribution. The coefficient of variation of the mixing distribution is often interpreted as a measure of the degree of heterogeneity. For the gamma distribution, CV reduces to

$$CV = \frac{SD[\lambda_0]}{E[\lambda_0]} = \frac{\sqrt{\frac{r}{\alpha^2}}}{\frac{r}{\alpha}} = \frac{1}{\sqrt{r}},$$

giving the r parameter the interpretation that a low (high) r value suggests a relatively more (less) heterogeneous population.

Control Variables

Table 6 provides the variable definitions, means, and standard deviations for the full model specification. Key model

TABLE 6
Variable Definitions

Variable	Definition	M (SD)
Really new product-market	1 for a really new product-market, 0 otherwise	.28 (.45)
Market pioneer	1 for the market's first entrant, 0 otherwise	.35 (.48)
Early follower	1 for each entrant in the first year of entry after the pioneer, 0 otherwise	.65 (.48)
Really new product-market × market pioneer	1 for a market pioneers in a really new product-market, 0 otherwise	.08 (.28)
Market pioneer × ln pioneer lead time	For the market pioneer, the natural logarithm of its lead time years over the second entrant plus 1.0, 0 otherwise	.40 (.61)
Early follower × ln pioneer lead time	For early followers, the natural logarithm of their delay after the pioneer's entry plus 1.0, 0 otherwise	.69 (.62)
Early follower × (ln pioneer lead time)^2	The squared value of the preceding interaction term	.86 (1.11)
Ln firm size	The natural logarithm of the firm's tangible asset size divided by its industry average	–.24 (1.82)
Incumbent	1 if the entering firm was in existence in period (t – 1), 0 if otherwise	.67 (.42)
Ln calendar year of market inception	The natural logarithm of the calendar year of market inception minus 1900	4.34 (.12)
Economic decline$_t$	1 if annual gross domestic product declined, 0 if otherwise	.18 (.38)
N_t	Normalized firm density at time t equals the number of competitors in the market divided by the average number over the ten-year period after the pioneer's monopoly ends	1.81 (1.85)
N_t^2	The squared value of the normalized firm density at time t divided by 100	.07 (.18)
N_f	The normalized firm density at the time of market entry	.38 (.31)
Chemicals	1 for a chemical product (SIC = 28), 0 otherwise	.07 (.26)
Electronics	1 for an electronic product (SIC = 36), 0 otherwise	.16 (.37)
Machinery	1 for a machinery product (SIC = 350–356), 0 otherwise	.20 (.40)
Instruments	1 for an instrument product (SIC = 38), 0 otherwise	.17 (.38)

Notes: For ln firm size, the industry average is defined at the two-digit SIC code. Dividing firm density by the average firm density of the market for ten years after the monopoly period normalizes the firm density variables. The means and standard deviations are based on 750 observations For the time-varying variables (Economic decline$_t$, N_t, and N_t^2), we measure the statistics at the last observation period of the sample.

specification variables we used in the hypothesis testing are the really new product-market dummy variable, a market pioneer dummy variable, and a really new product-market × market pioneer interaction term. Next, we discuss all the control variables.

Conventional economic variables. Geroski's (1995) literature review, suggests that the likelihood of survival is greater for larger firms than for smaller firms. The *Thomas Register* includes the tangible asset size for each firm. Because the average tangible asset size varies over industries and has changed over time, we compare entrant firm size at the time of market entry with the average industry firm size and categorize the entrants as larger versus smaller firms relative to the industry average. We define the industry at the two-digit SIC code level.

Incumbents are likely to have higher survival rates than start-up ventures because of their prior experience, market knowledge, brand equity, and customer relationships. In line with the work of Agarwal and Gort (1996), we define incumbents as firms in existence in year (t – 1).

In recent decades, increasing international competition and shorter new product development cycles have led to an increase in competitive intensity (Griffin 1997, p. 450). If competitive intensity has increased from the 1960s to the 1990s, hazard rates should be positively influenced by the calendar year of market inception.

We controlled for abnormally high hazard rates in the years of economic decline with a time-varying dummy variable. The dummy variable equals one in years when there is a decline in the annual gross domestic product growth rate. Audretsch (1991) finds that the ten-year firm survival rates vary considerably across different industries. Thus, the hazard rate model also includes four industry-specific dummy variables.

Population ecology variables. Population ecologists predict two contrasting effects of "density dependence" on firm hazard rates: legitimation and competition (Hannan and Carroll 1992; Hannan and Freeman 1989). Legitimation of a product innovation increases as the number of firms in the new product-market increases, thus resulting in the decline of hazard rates of all firms. However, there is a ceiling on the impact of the legitimacy process. At higher levels of firm density, legitimation reaches a saturation level, and the competitive effect of firm density dominates, resulting in higher hazard rates for all firms. Overall, the population ecology theory predicts a U-shaped relationship between firm density and hazard rate in new product-markets. Thus, the hazard rate model includes a firm density and a firm density–squared term.[4]

Density dependence theory also predicts a persisting competitive effect of firm density at the time of a firm's entry into a product-market. Carroll and Hannan (1989) attribute this effect to two sources: resource scarcity and tight niche packing. They contend that intense competition at the time of market entry creates the condition of resource scarcity. With scarce resources, market entrants that find it difficult to move to full-scale operation face strong selection pressures. Carroll and Hannan argue that another consequence of intense competition at the time of market entry is tight niche packing. Because market entrants seldom compete head-to-head with market incumbents, they are forced to pursue marginal target-market opportunities. Carroll and Hannan (1989), Carroll and colleagues (1996), and Banbury and Mitchell (1995) provide empirical evidence in support of the adverse impact of firm density at the time of market entry on the probability of firm survival.

Strategic variables. Because increasing pioneer lead time tends to strengthen first-mover advantages, we include pioneer lead time in the hazard rate model. Increasing pioneer lead time should decrease a market pioneer's hazard rate. For early followers, increasing pioneer lead time can increase their hazard rates because pioneers gain strength with longer lead time. However, a short delay can help them survive because in the first year or two of the market's evolution, decreased market and technological uncertainty can yield substantial benefits. In summary, we predict a U-shaped relationship for early follower hazard rates and pioneer lead time (Robinson and Min 2002). We use the natural logarithm of pioneer lead time to adjust for diminishing marginal returns.

We now present the full specification of the hazard rate's time-varying covariate function. For ease of presentation, deleting the i subscripts in the right-hand side of the expression for $\phi[X_i(t_i)]$ simplifies Equation 5's covariate function.

$$
\begin{aligned}
(14)\quad \phi[X_i(t_i)] = \mathrm{Exp}[&\beta_0 + \beta_1 \text{ (really new product-market)} \\
&+ \beta_2 \text{ (market pioneer)} \\
&+ \beta_3 \text{ (really new product-market} \\
&\times \text{market pioneer)} \\
&+ \beta_4 \text{ (market pioneer} \times \text{ln pioneer lead time)} \\
&+ \beta_5 \text{ (early follower} \times \text{ln pioneer lead time)} \\
&+ \beta_6 \text{ (early follower} \times \text{ln pioneer lead time}^2) \\
&+ \beta_7 \text{ (ln firm size)} + \beta_8 \text{ (incumbent)} \\
&+ \beta_9 \text{ (ln calendar year of market inception)} \\
&+ \beta_{10} \text{ (economic decline}_t) \\
&+ \theta_1 N_t + \theta_2 N_t^2 + \theta_3 N_f \\
&+ \lambda_s \sum_{s=1}^{S} \text{Industry}_s],
\end{aligned}
$$

where N_t is the normalized firm density at time t, N_t^2 is the square of the normalized firm density at time t, and N_f is the normalized firm density at the time of firm i's entry into the market.

In Equation 14, the really new product-market variable equals one for a really new product-market and zero for an incremental innovation product-market. The market pioneer dummy variable equals one for the first entrant and zero for an early follower. In Equation 14, the derivative of the logit of $\phi[X_i(t_i)]$ with respect to really new product-market is $\beta_1 + \beta_3$ (market pioneer). Thus, the estimated value of $\beta_1 + \beta_3$ tests H_1 (i.e., the effect of really new versus incrementally new products on the hazard rate of a market pioneer). If $\beta_1 + \beta_3$ is positive, pioneers have a higher exit rate for really new versus incrementally new products. For H_2, if β_1 is positive, early followers have a higher hazard rate in really new versus incrementally new product-markets.

To test H_3 (i.e., the effect of order of entry on survival risk in really new product-markets), we need to consider the differential impact of the order of entry (i.e., market pioneer versus early follower) on the hazard rate in Equation 14. Subtracting the derivative of $\phi[X_i(t_i)]$ with respect to early follower from the derivative of $\phi[X_i(t_i)]$ with respect to the market pioneer, we find that to test H_3, we need to consider the sign and significance of ($[\beta_2 + \beta_3 + \beta_4 \ln \text{lead time}] - [\beta_5 \ln \text{lead time} + \beta_6 \ln \text{lead time}^2]$). If this estimate is positive and significant, there is support for H_{3a}, indicating that the survival risk for market pioneers versus early followers is higher for really new products. Similarly, the estimate of ($[\beta_2 + \beta_4 \ln \text{lead time}] - [\beta_5 \ln \text{lead time} + \beta_6 \ln \text{lead time}^2]$) tests H_4. If this estimate is negative and significant, there is support for H_{4a}, indicating that pioneers experience lower hazard rates for incremental innovations than early followers.

Results

Table 7 displays the hazard rate model results. We based statistical significance on conservative two-tailed tests. In H_1, we predicted that market pioneers have a greater survival risk in really new product-markets. At the bottom of Table 7, market pioneers experience significantly greater survival risk for really new versus incrementally new products. Specifically, the survival risk of market pioneers is 189% ([exp(1.06) – 1] × 100%) greater for really new products than for incrementally new products. Recall that there is almost a threefold difference in the 12-year pioneer sur-

---

[4] To adjust for differences across product-markets, we normalize firm density by dividing it by the average number of competitors over the ten-year period following the end of the pioneer's monopoly.

TABLE 7
Hazard Rate Analysis (n = 750)

Variable		Estimate	t-Statistic
α		2651.40	.98
r		1.35	2.43**
D2		.63	2.88***
D3		.89	3.80***
D4		.99	3.78***
D5		1.00	3.39***
D6		1.07	3.30***
D7		1.15	3.25***
D8		1.40	3.76***
D9		1.60	4.06***
D10		1.65	3.86***
D11		1.67	3.61***
D12+		1.80	3.54***
Really new product-market	(β_1)	.20	1.20
Market pioneer	(β_2)	–1.53	–2.56***
Really new product-market × market pioneer	(β_3)	.86	2.90***
Market pioneer × ln pioneer lead time	(β_4)	–.34	–1.50
Early follower × ln pioneer lead time	(β_5)	–2 23	–2.52**
Early follower × (ln pioneer lead time)^2	(β_6)	.81	2.36**
Ln firm size	(β_7)	–.05	–1.47
Incumbent	(β_8)	–.34	–2.31**
Ln calendar year of market inception	(β_9)	1.31	5.14***
Economic decline$_t$	(β_{10})	.46	4.40***
N_t	(θ_1)	–.44	–2.79***
N_t^2	(θ_2)	6.83	2.83***
N_f	(θ_3)	1.05	4.19***
Chemicals	(γ_1)	–.04	–.18
Electronics	(γ_2)	.05	.29
Machinery	(γ_3)	.09	53
Instruments	(γ_4)	.21	1.22
Log-likelihood		–2081.5	
Test of Hypotheses			
H_1 ($\beta_1 + \beta_3$)		1.06	4.18***
H_2 (β_1)		.20	1.20
H_3 ($\beta_2 + \beta_3 + \beta_4$ ln lead time – β_5 ln lead time – β_6 ln lead time^2)		.43	1.77*
H_4 ($\beta_2 + \beta_4$ ln lead time – β_5 ln lead time – β_6 ln lead time^2)		–.44	–2.23**

*$p <$ 10.
**$p <$ 05.
***$p < .01$.

vival rate for really new versus incrementally new products (23% versus 61%). This indicates that for market pioneers, the challenge of resolving market and technological uncertainties in really new product-markets often outweighs the benefits of first-mover advantages.

In H_2, we predicted a greater survival risk for early followers in really new product-markets. In Table 7, the estimate has the expected sign, but it is not statistically significant. This result is consistent with the descriptive statistics: 12-year early follower survival rates are 38% for really new products versus 39% for incrementally new products. These results indicate that the typical 2- to 3-year delay in entering a really new product-market helps an early follower resolve key market and technological uncertainties.

For really new products, in H_{3a}, we predicted that market pioneers have a greater survival risk than early followers. At the bottom of Table 7, the difference is tested at the average pioneer lead time level of 1.98 years for the really new product sample. It is statistically significant at the 10% level. This difference is also managerially significant with a 54% ([exp(.43) – 1] × 100%) greater survival risk than early followers. This indicates that in really new product-markets, the challenge of managing the market and technological uncertainties often outweigh the benefits of first-mover advantages.

In contrast, H_{4a} reveals that for incremental innovations, market pioneers have a substantially higher chance of survival than early followers. These market pioneers experience a 36% (= [exp(–.44) –1] × 100%) lower survival risk than early followers, which is significant at the 5% level. This supports the prediction that for incrementally new products, market pioneers have a greater chance of survival.

Turning now to the coefficient estimates of the control variables, we observe that they all have the expected signs. The impact of lead time on the pioneer survival chance has the expected negative sign and is statistically significant at the 13% level. This indicates that increasing pioneering lead time reduces a market pioneer's survival risk.

The results support a U-shaped relationship between an early follower's hazard rate and a pioneer's lead time. Thus, a short delay appears to help an early follower, but an additional delay increases an early follower's survival risk. Incumbency significantly decreases survival risk. Furthermore, the calendar year of market inception has the expected positive sign and is statistically significant. This indicates increasing survival risk in markets started in recent years. There is also a significantly greater survival risk in years of economic decline.

The estimates of θ_1 and θ_2 show the impact of firm density on a firm's hazard rate. As expected, the estimates for θ_1 and θ_2 are negative and positive, respectively. Both estimates are statistically significant at the 1% level. This U-shaped relationship indicates that an initial increase in the number of firms increases its legitimation, thus lowering firm hazard rates. At higher levels of firm density, legitimation reaches a saturation level, and the competitive effect dominates with an associated adverse impact on firm survival rates.

As predicted, the estimate of θ_3 is negative and statistically significant at the 1% level. This indicates that an increase in firm density at the time of market entry elevates a firm's hazard rate. Population ecologists attribute this result to the liability of resource scarcity and to tight niche packing, which makes it difficult for an entrant to find profitable positioning opportunities.

Robustness of Results

Because theory does not provide clear guidance on the functional relationship between the degree of product innovation newness and survival, we consider two additional functional forms. The first maintains the dichotomy we used previously but considers a product really new only when every expert classifies it as really new. This differs from our measure, which is based on the majority of expert opinion. When every expert classifies a product as really new, our sample has 30 "really" really new products.

The second functional form uses a continuous measure of product innovation newness. This follows the research of Chandy and Tellis (1998, 2000), whose measure of radical product innovations is continuous. Our continuous measure equals the difference between the expert's rating and the expert's average benchmark rating. We average these differences across the experts who evaluated each of the 264 new markets.

Table 8 provides the hypothesis-testing results for these alternative functional forms. Model 1 tests the dichotomy between "really" really new products and incremental innovations. (We exclude the 36 really new product-markets that do not have uniform expert support from the data analysis.) Model 2 uses the continuous measure of new product newness.

Overall, the hypothesis-testing results are consistent with all three measures of product innovation newness. This is because the results consistently indicate that market pioneers have higher hazard rates in markets that were pioneered with a really new product but have lower hazard rates in markets that were pioneered with an incremental innovation. In addition, there is no material difference between early follower hazard rates in really new versus incremental innovation markets.

Dynamic Survival Risk Insights

Although our hazard rate model yields important survival risk insights, it does not pinpoint how survival risk changes over time. We develop two types of dynamic insights. The empirical hazard rate equals the number of firms that failed in year t divided by the total number of firms that were at risk in year (t – 1). Because empirical hazard rates do not partial out the impact of additional explanatory variables or correct for unobserved heterogeneity, we also provide predicted hazard rates.

The predicted hazard rates are based on four hazard rate models. We estimate the four models for market pioneers (n = 66) and early followers (n = 146) in really new product-markets and for market pioneers (n = 198) and early followers (n = 340) in incrementally new markets. The full model specification includes the additional explanatory variables in Table 7. Because the incremental innovations sample is larger, we used annual dummy variables for the first 12 years of commercialization versus 10 years for the smaller really new product-markets sample.

For really new product-markets, Panels A and B in Figure 1 display the empirical and predicted hazard rate plots, respectively. These figures illustrate how hazard rates for pioneers and early followers are initially similar but then diverge. The divergence arises after the sixth year of commercialization. Note that the tenth year includes year ten plus all subsequent years. This pattern can arise when it takes a few years for the market pioneer to learn whether its initial investments are on target. During the wait, the market pioneer will fight to survive, but after it is clear that its initial investments are off target and cannot be salvaged, it is time to cut its losses and move on to better opportunities.

For incrementally new product-markets, Panels A and B in Figure 2 show that market pioneers have the greatest survival advantages in the first six to eight years of commercialization. This can arise when market pioneers benefit from their temporary monopoly in the first two or three years of commercialization. In addition, because first-mover advantages often dissipate over time (Robinson 1988), it is not surprising that hazard rate differences narrow with the passing of time. Even so, market pioneers have consistently lower hazard rates than early followers.

Discussion

When a market pioneer starts a new market with a really new product, survival is a major challenge. We base this conclusion on 66 really new product-markets, in which only 23% of the market pioneers survived for at least 12 years. In

TABLE 8
Robustness of Results

Variable	Model 1 (n = 645)	Model 2 (n = 750)
α	3337.21 (.79)	1908.30 (1.10)
r	1.22 (2.07)**	1.49 (2.46)**
D2	.64 (2.67)***	.61 (2.83)***
D3	.91 (3.41)***	.86 (3.76)***
D4	1.13 (3.73)***	.95 (3.78)***
D5	1.07 (3.05)***	.95 (3.39)***
D6	1.15 (2.96)***	1.02 (3.31)***
D7	1.40 (3.28)***	1.09 (3.28)***
D8	1.37 (2.93)***	1.34 (3.85)***
D9	1.68 (3.45)***	1.53 (4.17)***
D10	1.77 (3.36)***	1.58 (3.96)***
D11	1.87 (3.30)***	1.60 (3.70)***
D12+	2.04 (3.25)***	1.71 (3.67)***
Innovativeness		.02 (.53)
"Really" really new product-market	.26 (1.18)	
Market pioneer	–1 53 (–2.32)**	–1.01 (–1.76)*
Innovativeness × market pioneer		.30 (3.98)***
"Really" really new product-market × market pioneer	1.10 (2.57)***	
Market pioneer × ln pioneer lead time	–.41 (–1.56)	–.26 (–1.18)
Early follower × ln pioneer lead time	–2.28 (–2.31)**	–2.41 (–2.79)***
Early follower × (ln pioneer lead time)^2	.81 (2.15)**	.86 (2.58)***
Ln firm size	–.04 (–1.14)	–.03 (–.93)
Incumbent	–.30 (–1.81)*	–.34 (–2.43)**
Ln calendar year of market inception	1.39 (4.45)***	1.26 (5.36)***
Economic decline$_t$	.45 (3.93)***	.46 (4.40)***
N_t	–.55 (–2.89)***	–.42 (–2.95)***
N_t^2	7.91 (2.54)**	6.35 (3.14)***
N_f	1.18 (4.16)***	1.01 (4.16)***
Chemicals	.36 (1.17)	–.01 (–.06)
Electronics	.01 (.04)	.05 (.29)
Machinery	.07 (.37)	.10 (.65)
Instruments	.34 (1.62)	.22 (1.32)
Log-likelihood	–1780.9	–2077.6
Hypotheses		
H_1	1.36 (3.58)***	.32 (4.98)***
H_2	.26 (1.18)	.02 (.53)
H_3	.64 (1.73)*	.52 (2.40)**
H_4	–.47 (–2.04)**	–.43 (–4.61)***

$^{*}p < .10$.
$^{**}p < .05$
$^{***}p < .01$.
Notes: To test H_3 and H_4 for Model 1 specification, we evaluated the derivative difference between market pioneers' and early followers' hazard rates with the average values of "innovativeness" from 66 really new product-markets (=.68) and 198 incrementally new product-markets (=–2.53), respectively.

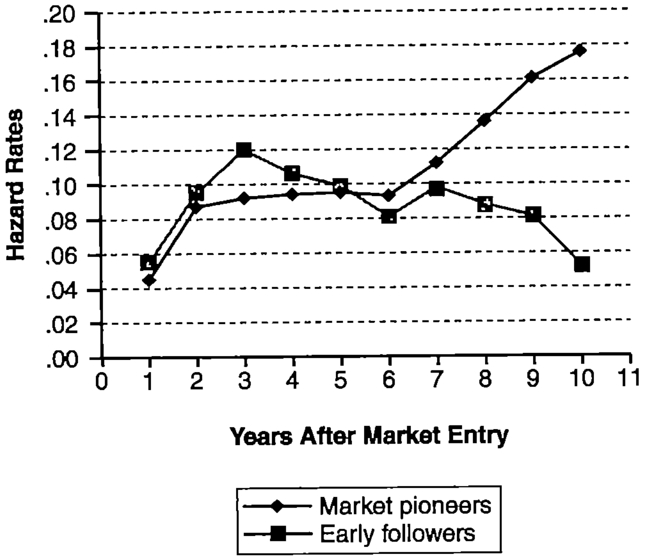
contrast, survival is much easier when the pioneer starts a new market with an incremental innovation. We base this conclusion on 198 incrementally new product-markets, in which 61% of the market pioneers survived for at least 12 years.

These results indicate that market pioneers must overcome greater resistance in building primary demand for really new products. Conversely, the market pioneer's temporary monopoly can help the pioneer strengthen its first-mover advantages for incrementally new products that represent extensions of existing technologies and, in general, do not require consumer learning.

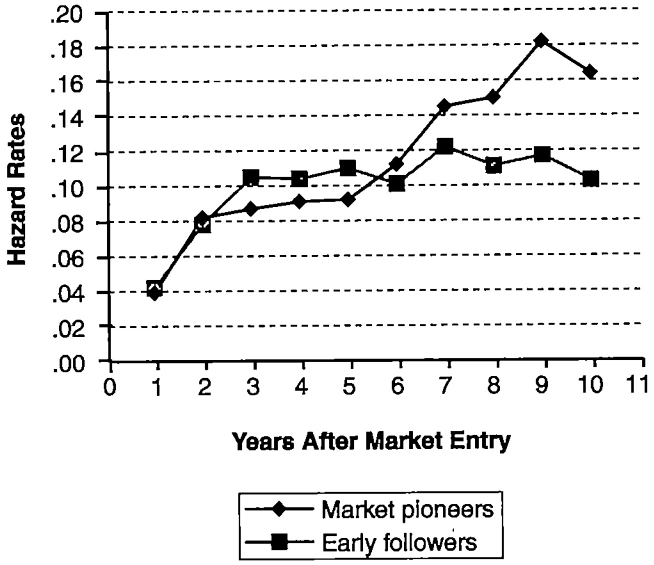
In really new product-markets, is this high degree of uncertainty also fatal for early followers? In our sample, most of the early followers trailed the market pioneer by three years or less. It is surprising that even with this relatively short delay, there is not a meaningful survival rate difference across really new and incrementally new product-markets. For example, 12-year early follower survival rates in really new product-markets are 38% versus 39% in incrementally new product-markets. This suggests that in really new product-markets, early followers can benefit from the prior efforts of the pioneer in building primary demand. They can also benefit from the pioneer's experience in addressing customer needs and concerns, which can become public knowledge through professional and trade journals (Gort and Klepper 1982). Finally, early followers are more likely to leapfrog the pioneer in really

FIGURE 1
Hazard Rates In Really New Product-Markets

A: Empirical Hazard Rates



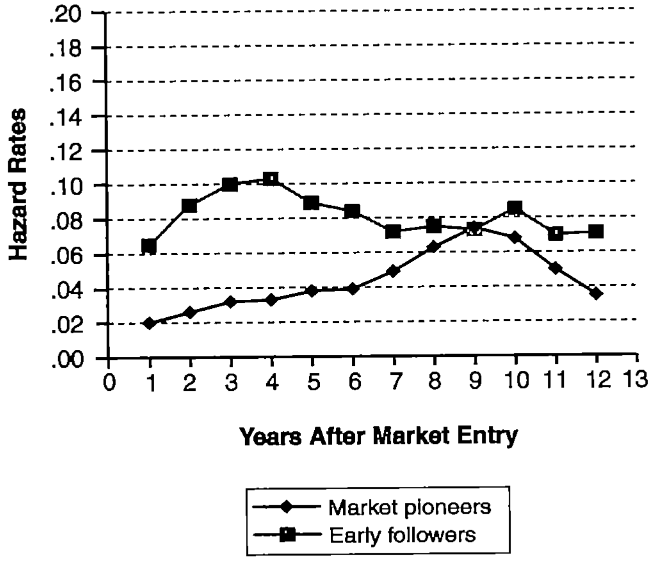
B: Predicted Hazard Rates



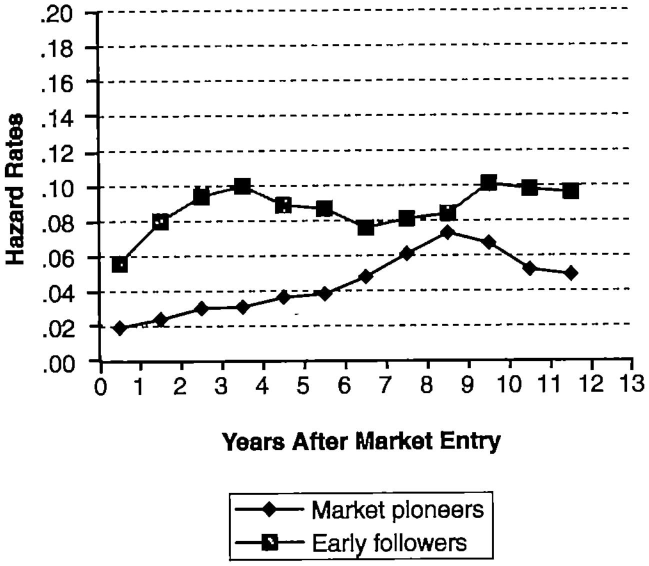
Notes: We use three-year moving averages except for the first year. For predicted hazard rates, we assume that all the firms do not face economic decline. The tenth year hazard rate in the predicted hazard rate function indicates the hazard rate in the tenth year and thereafter.

FIGURE 2
Hazard Rates In Incremental Innovation Product-Markets

A: Empirical Hazard Rates



B: Predicted Hazard Rates



Notes: We use three-year moving averages except for the first year. For predicted hazard rates, we assume that all the firms do not face economic decline. The 12th year hazard rate in the predicted hazard rate function indicates the hazard rate in the 12th year and thereafter.

new product-markets that often undergo significant product improvements during the emergent stages (Christensen, Suarez, and Utterback 1998).

In addition to highlighting the survival risks of pioneering really new product-markets, these empirical results help resolve an ongoing controversy in the order-of-market-entry literature. The controversy is whether the first to market is the first to fail. In really new product-markets, the hazard rate modeling plots show that greater market pioneer survival risks tend to surface after six or more years in the market. This can arise when it takes a few years for a market pioneer to determine whether its initial investments, which were made in the face of great uncertainty, are on target.

For incrementally new product-markets, the hazard rate modeling plots show that market pioneers have lower survival risks than early followers. Although these survival risk differences narrow over time, market pioneers have consistently lower survival risk. This can be explained by pioneers' initial monopoly, which increases survival rates in the first two to three years of commercialization. Furthermore, pioneer first-mover advantages are often strongest in the early years of a market's evolution, and these advantages help protect the pioneer from outright failure.

In conclusion, market pioneers are often the first to fail in really new product-markets. However, this is not true in incrementally new markets, in which market pioneers have consistently lower survival risks than early followers.

Research Limitations and Further Research

We now discuss four research limitations and the associated future research topics. First, our sample is composed only of industrial goods markets. Different results could surface in different types of markets. For example, incremental innovations in high-technology markets may not generate meaningful pioneer survival benefits relative to early followers. This is because a pioneer's survival can be threatened by a higher degree of technological uncertainty.

Second, our sample contains only 66 really new product-markets. Although this sample size compares favorably with other survival studies, a larger sample size would yield more precise insights into market pioneer survival risks and how these survival risks change over time.

Third, our product innovation newness measure is applied decades after the new markets started. Major investments in new products arise before product launch, perhaps even at the prototype development stage. Thus, during the product development process, managers need a different type of measure to distinguish really new from incrementally new products.

Fourth, optimal entry timing requires greater insights into fast followers versus slow followers. For example, how long should an early follower wait before attempting to enter a new market? Although the *Thomas Register* data can address this research topic, it will require an expanded sample size and an objective way to distinguish fast followers from slow followers.

Managerial Implications

Pioneering a new market is often considered a high-risk investment with high potential returns. If this is indeed the case, we expect that the highest returns will surface in the riskiest markets. However, at least from the standpoint of survival, this is not the case. This is because market pioneers have the lowest survival rates in really new product-markets, which carry the highest risk. Given this pattern of results, why would a firm attempt to pioneer a really new product-market?

A key reason is that if lower survival rates are compensated with higher profits, pioneers might still be better off.[5] Higher profits can arise because really new product-markets tend to be larger with higher profit margins (Mansfield and Wagner 1975). This indicates that first-mover advantages can be especially strong for the few pioneering survivors. For example, according to the *Thomas Register* data, in 1981, Hewlett-Packard, DatagraphiX, and IBM were the first laser printer entrants. Hewlett-Packard is a dominant survivor, and DatagraphiX exited in 1993. Thus, for really new products, though the risks of pioneering are high, so are the payoffs for the pioneers that do survive.

Lower survival risks for early followers in really new product-markets suggest that early following can also be a profitable strategy. This is because the early followers not only face a lower survival risk than market pioneers but also can benefit from the market's larger size and higher profit margins. These results are consistent with the work of Olleros (1986, p. 11), who contends that breakthrough innovations "are bound to be rather crude, costly, and unreliable when they first appear in the market." This provides an opportunity for early followers to leapfrog pioneers in really new product-markets.

The risk-and-return profile in incremental innovation markets is more encouraging for market pioneers. Incremental innovation markets tend to have relatively low risk and higher pioneer survival rates. Incremental innovation markets are also likely to yield sustainable market share advantages for the pioneer. This is because research on surviving entrants points to pioneer market share rewards (see, e.g., Kalyanaram, Robinson, and Urban's [1995] survey and VanderWerf and Mahon's [1997] meta-analysis). With the majority of new markets being pioneered by incremental innovations, it is likely that these market share reward studies are dominated by incremental innovation markets. If so, pioneers of incremental innovation markets should benefit from higher survival rates and sustainable market share advantages.

---

[5] Although profit data would yield valuable research insights, they are not provided in the *Thomas Register*. Profit data are available in PIMS, but the business unit and product category are both disguised. With disguised product categories, really new and incrementally new product-markets cannot be identified. Thus, we do not know of any profit data that can be linked to really new versus incrementally new product-markets.

APPENDIX
List of Product-Markets

Really New Products (Year of Market Inception, Degree of Innovativeness)

Chemical Products

Bromide lead (1974, .1), carbon monoxide gas (1962, –.3), cellulose propionate (1963, .1), dielectric gas (1976, .1), gelling agents (1966, 1.3), indium compounds (1980, –.1), lanthanides (1985, .2), lithium acetate (1975, –.2), lithium niobate (1985, .5), magnesium perchlorate (1971, .0), methyl laurate (1964, –.1), microbiocides (1982, .4), and nickel carbonyl (1973, –0.4)

Materials and Fabricated Metal

Free-machining steel (1971, 1.3), parylene (1987, .6), polycarbonate film (1981, 1.2), polyimide film (1981, 1.7), shape memory alloys (1988, 1.6), solar energy reflector sheets (1982, 1.6), and solar energy storage rods (1982, 2.0)

Machinery and Computer Equipments

Blood gas monitors (1989, .6), clean room hoods (1967, .3), computer terminals (1971, 1.8), data display terminals (1971, 1.4), digital comparators (1967, .6), graphite machinery (1986, –.2), handheld scanners (1990, .6), positioning machinery (1964, .6), protocol translators (1986, 1.4), sealing robots (1986, 1.1), steam turbogenerators (1982, .6), transaction processing systems (1987, 1.9), and ultrasonic drilling equipment (1971, 1.4)

Electronic and Electrical Equipment

Electroluminescent lighting (1982, .0), electron beam guns (1965, .7), ion beam guns (1965, 1.0), logic systems (1967, 1.9), memory cards (1991, .6), phase detectors (1970, .7), photoelectric timers (1972, 1.1), silicon detectors (1973, 1.0), silicon rectifier converters (1968, .1), videodisc recorders (1985, 1.6), videotape recorders (1973, 1.6), and voice activated microphones (1992, 1.6)

Measuring, Analyzing, and Controlling Equipments

Blood pressure monitors (1986, –.3), blood pressure transducers (1988, .6), bomb detectors (1967, .0), bubble presence detectors (1962, –.1), computer screen projectors (1986, .6), ultraviolet detectors (1971, 1.1), ellipsometers (1987, .4), fiber optic attenuators (1985, .1), fiber optic controls (1986, .6), frequency translators (1986, .6), germanium lenses (1963, .5), microfilm reader/printers (1971, .2), microscopy cameras (1995, 1.1), navigating systems (1967, 1.0), osmometers (1964, .4), video microscopes (1986, .0), and videotape motion analysis systems (1983, 1.2)

Other Industries

Bar code computer software (1987, 1.0), EMC testing computer software (1995, .6), emulation computer software (1990, .6), and travel computer software (1995, .6)

Incrementally New Products (Year of Market Inception, Degree of Innovativeness)

Chemical Products

Chrome dyes (1973, –1.7), gallium oxide (1987, –1.7), laser printing inks (1987, .1), lead silico fluoride (1976, –2.2), lead titanate (1964, –2.8), magnesium zirconate (1986, –2.8), manganese benzoate (1979, –1.3), manganous sulfate (1964, –3.4), methyl mercaptan (1973, –3.2), molybdenum disilicide (1979, –2 2), nickel iodide (1975, –3.3), and nickel sulfamate (1969, –3.2)

Materials and Fabricated Metal

Access, animal doors (1972, –3.7), adjustable nipples (1972, –2.4), air duct joints (1962, –3 3), air logic valves (1986, –.3), asphalt concrete (1963, –2.3), bearing pads (1987, –3.4), boiler chimneys (1976, –3.7), carbide seals (1971, –1.4), chemical resistant seals (1982, –1.4), chromatography packing (1987, –3.4), conveyor belt lugs (1970, –1.3), diamond discs (1977, –1.0), disc dispensers (1979, –2.7), drain orifices (1975, –1.0), fiberglass reinforced plastic windows (1986, –2.4), fiberglass washers (1971, –1.4), frostproof valves (1967, –2.4), ground, wire, plastic mouldings (1964, –2.3), high-pressure seals (1982, –1.4), hopper outlets (1964, –1.0), kiln vents (1986, –4.4), laminating rollers (1971, –1.4), laser mirrors (1966, –1.7), low-temperature belting (1966, –2.4), magnetic ceramics (1967, –.4), master mirrors (1964, –2.4), microminiature tubes (1987, –1.4), molten aluminum (1966, –2 0), mud, drilling lubricants (1981, –2.0), newsstands (1986, –3 7), niobium foil (1969, –.3), noise mufflers (1973, –2.7), passivating equipment (1986, –1.4), plastic netting (1966, –2.3), plastic sheaths (1987, –2.4), plunger tips (1986, –3.4), prefabricated chimneys (1976, –3.0), refractory foil (1967, –1), reprocessed oils (1981, –2.0), sapphire orifices (1972, –1.7), shock-absorbing dampers (1962, –2.0), silica curtains (1977, –.7), silo liners (1966, –.7), soil shifting screens (1964, –3.4), solar collector drain-down valves (1982, –3.4), stainless steel bumpers (1987, –2.4), steel foil (1967, –1.0), tantalum plates (1986, –1.4), transistor insulators (1962, –.3), urethane gaskets (1971, –.7), vise grips (1972, –3.0), and zirconium copper (1962, –2.3)

Machinery and Computer Equipments

Adjustable gears (1968, –2.2), air movers (1975, –3.0), aluminum rollers (1987, –3.9), boat coolers (1976, –3.7), boat loaders (1973, –3.5), bottle sorters (1971, –2 9), bullet molds (1975, –3 4), canners coolers (1962, –3.3), carbide rings (1971, –3.4), chamfering heads (1973, –3.2), commutator turning lathes (1980, –2.2), conveyor belt guides (1970, –2.9), copy-turning machinery (1970, –2.0), cosmetic molds (1964, –2.7), cross-drilling (1965, –3.9), cutterhead dredges (1968, –2.7), diamond-machining machinery (1975, –2.2), egg carton machinery (1964, –3.3), flash-drying mills (1973, –3.8), fluid

APPENDIX
Continued

separators (1971, –2.9), fogging machinery (1975, –3.4), forced draft, fan mufflers (1966, –2.8), friction plates (1987, –3.4), furniture molds (1964, –3.4), gas liquefiers (1973, –2.0), glass crushers (1972, –3.4), ice-dicing machinery (1976, –3.7), jig grinders (1973, –3.7), lateral collectors and distributors (1981, –3.9), log chippers (1972, –3.8), lowering devices (1972, –4.0), magnetic tape erasers (1962, –2.3), mask washers (1975, –3.9), mixing mills (1973, –3.9), molding, shoe counter machinery (1964, –3.5), motor assembly machinery (1975, –2.4), mud jacking (1973, –2.8), oil-filtering and transfer equipment (1982, –3.4), oil recovery systems (1975, –2.5), opening machinery (1973, –3.7), package sorters (1964, –2.4), pellet coolers (1962, –3.9), pickle-making machinery (1986, –4.0), poultry chillers (1962, –4.2), pug, agricultural chemicals mills (1972, –3.2), radio frequency identification scanners (1989, –2.4), roll-forming machinery (1970, –3.2), seaport tractors (1982, –2.9), shot sleeves (1986, –4.4), sign rotators (1967, –2.4), snow-melting mats (1966, –3.8), solid/liquid mixers (1966, –3.8), spherical grinding (1962, –2.2), steel belt coolers (1962, –2.7), thermal liquid coolers (1982, –3.2), thermoforming machinery (1970, –1.4), torque multipliers (1979, –1.9), tray loaders (1975, –4.0), tubing pumps (1982, –2.4), welding nozzles (1987, –2.6), wet-grinding mills (1964, –3.0), and wood-chipping machinery (1968, –2.5)

Electronic and Electrical Equipment

Bar code lamps (1985, –.1), boat lights (1987, –3.6), circuit board retainers (1987, –1.4), D C. micro motors (1987, –1.0), explosion detectors (1964, –1.7), explosion-proof lamps (1969, –4.1), fiber optic lamps (1991, –2.4), fiber optic links (1993, –2.4), humidity transmitters (1986, –.4), insect control systems (1971, –2.7), magnifying lamps (1964, –1.4), pilferage detectors (1964, –1.9), railroad car lights (1987, –4.1), relay-switching matrices (1973, –3.7), RF communication systems (1992, –3.4), robot guarding systems (1987, –1.6), solid state contactors (1967, –2.1), synchronous generators (1972, –2.3), thermal cutoffs (1975, –2.0), V-band communication systems (1990, –3.4), video modems (1992, –3.4), video modulators (1992, –2.4), VSAT communication systems (1993, –3.4), and xenon lamps (1966, –1.4)

Measuring, Analyzing, and Controlling Equipments

Absorptometers (1966, –1.8), water analyzers (1966, –2.1), aperture cards (1962, –2.6), backer rods (1986, –1.4), blood-processing equipment (1968, –.4), diesel fuel meters (1972, –3.5), distance-measuring instruments (1972, –2.1), electric sensors (1964, –1.6), hour meters (1967, –3.0), instant photo identification systems (1973, –.9), liquid medical waste absorbents (1993, –3.4), medical electronic filters (1992, –4.4), medical equipment hydraulic power units (1988, –3.4), medical waste disinfection systems (1992, –3.4), microscope objectives (1987, –1.6), multitesters (1985, –1.7), optical sensors (1964, –.9), photographic transparency illuminators (1987, –1.0), pneumatic to electric converters (1987, –.9), pressure compensators (1962, –2.5), protective masks (1979, –3.5), radio annunciators (1989, –2.4), roundness testers (1972, –3.4), scanning tunneling microscopes (1992, -2.4), static electricity detectors (1973, -2.5), and web deflection indicators (1985, –1.1)

Other Industries

Air-operated doors (1975, –2.5), battery pads (1986, –3.4), coin-operated lockers (1968, –3.2), credit card imprinters (1962, –2.2), disposable mops (1975, –2.4), doorway sealers (1966, –5.2), doorway-sealing equipment (1964, –5.2), hardboard letters and figures (1975, –5.5), lactose (1985, –4.2), letter shop (1977, –5.2), mitre fold construction (1973, –3.5), mobile comfort stations (1973, –2.5), numerical control tapes (1971, –2.4), overall trimmings (1986, –3.4), release film (1981, –3.2), steel tire cord (1975, –4.8), synthetic fibre cordage (1966, –3.8), transfer tapes (1964, –2.4), web strap trimmings (1986, –3.4), flatbed vehicles (1986, –3.6), skin markers (1966, –4.8), and sign kits (1968, –5.2)

Notes: The degree of innovativeness equals the average expert rating minus the benchmark rating for really new products that are well known. Although most values are positive for really new products and negative for incrementally new products, there are a few exceptions, which arise because the classification is based on the majority of expert opinion and not on the average expert rating.

REFERENCES

Agarwal, Rajshree (1997), "Survival of Firms over the Product Life Cycle," *Southern Economic Journal*, 63 (3), 571–84

——— and Barry L. Bayus (2002), "The Market Evolution and Sales Take-Off of Product Innovations," *Management Science*, 48 (August), 1024–1041.

——— and Michael Gort (1996), "The Evolution of Markets and Entry, Exit, and Survival of Firms," *Review of Economics and Statistics*, 78 (3), 489–98.

——— and ——— (2001), "First-Mover Advantage and the Speed of Competitive Entry," *Journal of Law and Economics*, 44 (April), 161–77.

Audretsch, David B. (1991), "New Firm Survival and the Technological Regime," *The Review of Economics and Statistics*, 3 (August), 441–50.

Banbury, C.M and W. Mitchell (1995), "The Effect of Introducing Important Incremental Innovations on Market Share and Business Survival," *Strategic Management Journal*, 16 (Special Issue), 161–82.

Brown, Christina and James M. Lattin (1994), "Investigating the Relationship Between Time in Market and Pioneering Advantage," *Management Science*, 40 (October), 1361–69.

Carpenter, Gregory S. and Kent Nakamoto (1989), "Consumer Preference Formation and Pioneering Advantage," *Journal of Marketing Research*, 26 (August), 285–98.

Carroll, Glenn R., Lyda S. Bigelow, Marc-David L. Seidel, and Lucia B. Tsai (1996), "The Fates of *De Novo* and *De Alio* Producers in the American Automobile Industry 1985-1981," *Strategic Management Journal*, 17 (Special Issue), 117–37.

——— and Michael T. Hannan (1989), "Density Delay in the Evolution of Organizational Populations: A Model and Five Empirical Tests," *Administrative Science Quarterly*, 34 (September), 411–30.

Chandy, Rajesh K. and Gerard J. Tellis (1998), "Organizing for Radical Product Innovation: The Overlooked Role of Willingness to Cannibalize," *Journal of Marketing Research*, 35 (November), 474–87.

——— and ——— (2000), "The Incumbent's Curse? Incumbency, Size, and Radical Product Innovation," *Journal of Marketing*, 64 (July), 1–17.

Christensen, Clayton M., Fernando F. Suarez, and James M. Utterback (1998), "Strategies for Survival in Fast-Changing Industries," *Management Science*, 44 (12), S207–S220.

Garcia, Rosanna and Roger Calantone (2002), "A Critical Look at Technological Innovation Typology and Innovation Terminology: A Literature Review," *Journal of Product Innovation Management*, 19 (March), 110–132.

Geroski, P.A. (1995), "What Do We Know About Entry," *International Journal of Industrial Organization*," 13 (4), 421–40.

Golder, Peter N. and Gerard J. Tellis (1993), "Pioneer Advantage: Marketing Logic or Marketing Legend?" *Journal of Marketing Research*, 30 (May), 158–70.

Gort, Michael and Stephen Klepper (1982), "Time Paths in Diffusion of Product Innovations," *Economic Journal*, 92 (367), 630–53.

Griffin, Abbie (1997), "PDMA Research on New Product Development Practices: Updating Trends and Benchmarking Best Practices," *Journal of Product Innovation Management*, 6 (November), 429–58.

Han, A. and J. Hausman (1990), "Flexible Parametric Estimation of Duration and Competing Risk Models," *Journal of Econometrics*, 5 (1), 1–28.

Hannan, M.T. and G.R. Carroll (1992), *Dynamics of Organizational Populations*. New York. Oxford University Press.

——— and J. Freeman (1989), *Organizational Ecology*. Cambridge, MA: Harvard University Press.

Heckman, J. and B. Singer (1984), "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52 (2), 271–320.

Helsen, Kristiaan and David C. Schmittlein (1993), "Analyzing Duration Times in Marketing: Evidence for Effectiveness of Hazard Rate Models," *Marketing Science*, 11 (Fall), 395–414.

Huff, Lenard C. and William T. Robinson (1994), "The Impact of Leadtime and Years in the Market on Pioneer Market Share Advantages," *Management Science*, 40 (October), 1370–77.

Kalyanaram, Gurumurthy, William T. Robinson, and Glen L. Urban (1995), "Order of Market Entry: Established Empirical Generalizations, Emerging Generalizations, and Future Research," *Marketing Science*, 14 (2), G212–G221.

Kerin, Roger A., P. Rajan Varadarajan, and Robert A. Peterson (1992), "First-Mover Advantage: A Synthesis, Conceptual Framework, and Research Propositions," *Journal of Marketing*, 56 (October), 33–52.

Lambkin, Mary and George S. Day (1989), "Evolutionary Processes in Competitive Markets: Beyond the Product Life Cycle," *Journal of Marketing*, 53 (July), 4–20.

Lancaster, T. (1979), "Econometric Methods for the Duration of Unemployment," *Econometrica*, 47 (4), 939–56.

——— (1990), *The Econometric Analysis of Transition Data*. Cambridge, UK: Cambridge University Press.

Lavin, Michael R. (1992), *Business Information: How to Find It, How to Use It*. Phoenix: Oryx Press.

Lieberman, Marvin B. and David B. Montgomery (1988), "First-Mover Advantages," *Strategic Management Journal*, 9 (5), 41–58.

Lilien, Gary L. and Eungsang Yoon (1990), "The Timing of Competitive Market Entry: An Exploratory Study of New Industrial Products," *Management Science*, 36 (5), 568–85.

Mansfield, Edwin and Samuel Wagner (1975), "Organizational and Strategic Factors Associated with Probabilities of Success in Industrial R&D," *Journal of Business*, 48 (April), 179–98.

Meyer, B.D. (1990), "Unemployment, Insurance and Unemployment Spells," *Econometrica*, 58 (4), 757–82.

Mitchell, Will (1991), "Dual Clocks: Entry Order Influences on Incumbent and Newcomer Market Share and Survival When Specialized Assets Retain Their Value," *Strategic Management Journal*, 12 (2), 85–100.

Olleros, Francisco-Javier (1986), "Emerging Industries and the Burnout of Pioneers," *Journal of Product Innovation Management*, 1 (1), 5–18.

Robinson, William T. (1988), "Sources of Market Pioneer Advantages: The Case of Industrial Goods Industries," *Journal of Marketing Research*, 25 (February), 87–94.

——— and Claes Fornell (1985), "Sources of Market Pioneer Advantages in Consumer Goods Industries," *Journal of Marketing Research*, 22 (August), 305–317.

——— and Sungwook Min (2002), "Is the First to Market the First to Fail? Empirical Evidence for Industrial Goods Businesses," *Journal of Marketing Research*, 34 (February), 120–28.

Schmalensee, Richard (1982), "Product Differentiation Advantages of Pioneering Brands," *American Economic Review*, 27 (June), 349–65.

Schmidt, Jeffery B. and Roger J. Calantone (1998), "Are Really New Product Development Projects Harder to Shut Down?" *Journal of Product Innovation Management*, 15 (2), 111–23.

Shepherd, Dean A. (1999), "Venture Capitalists' Assessment of New Venture Survival," *Management Science*, 45 (May), 621–32.

Song, X. Michael and Mitzi M. Montoya-Weiss (1998), "Critical Development Activities for Really New Versus Incremental Products," *Journal of Product Innovation Management*, 15 (2), 124–35.

Srinivasan, Raji, Gary L. Lilien, and Arvind Rangaswamy (2004), "First In, First Out? The Effects of Network Externalities on Pioneer Survival," *Journal of Marketing*, 68 (January), 41–58.

Szymanski, David M., Lisa C. Troy, and Sundar G. Bharadwaj (1995), "Order of Entry and Business Performance: An Empirical Synthesis and Reexamination," *Journal of Marketing*, 59 (October), 17–33.

Tellis, Gerard J. and Peter Golder (1996), "First to Market, First to Fail? Real Causes of Enduring Market Leadership," *Sloan Management Review*, 2 (Winter), 65–75.

Urban, Glen L., Theresa Carter, Steven Gaskin, and Zofia Mucha (1986), "Market Share Rewards to Pioneering Brands: An Empirical Analysis and Strategic Implications," *Management Science*, 32 (June), 645–59.

———, Bruce D. Weinberg, and John R. Hauser (1996), "Premarket Forecasting of Really-New Products," *Journal of Marketing*, 60 (January), 47–60.

VanderWerf, Pieter A. and John F. Mahon (1997), "Meta-Analysis of the Impact of Research Methods on Findings of First-Mover Advantage," *Management Science*, 43 (November), 1510–19.

Vanhuele, Marc, Marnik G. Dekimpe, Sunil Sharma, and Donald G. Morrison (1995), "Probability Models for Duration: The Data Don't Tell the Whole Story," *Organizational Behavior and Human Decision Processes*, 62 (April), 1–13.

Yamaguchi, K. (1991), *Event History Analysis: Applied Social Research Methods Series*, Vol. 28. Newbury Park, CA: Sage Publications.

Tülin Erdem, Joffre Swait, & Ana Valenzuela

Brands as Signals: A Cross-Country Validation Study

This article tests how well the information economics view of brand equity explains consumer brand choice in countries that represent different cultural dimensions. In this empirical analysis, the authors use survey and experimental data on orange juice and personal computers collected from respondents in Brazil, Germany, India, Japan, Spain, Turkey, and the United States. The results provide strong empirical evidence across countries for the role of brands as signals of product positions. In addition, the positive effect of brand credibility on choice is greater for consumers who rate high on either collectivism or uncertainty avoidance. Credible brands provide more value to collectivist consumers because such consumers perceive these brands as being of higher quality (i.e., reinforcing group identity). Credible brands provide more value to high-uncertainty-avoidance consumers because such brands have lower perceived risk and information costs.

A brand—understood to be "a name, term, sign, symbol, or design, or a combination of them, [that] is intended to identify the goods and services of one seller or a group of sellers and to differentiate them from those of competitors" (Kotler 1997, p. 443)—can potentially play many roles in consumer decision making and choice behavior. Underlying many of the brand effects is consumer uncertainty about product attributes and/or benefits. This uncertainty arises from the imperfect and asymmetric information state that characterizes most product markets (i.e., firms are more informed about their own products than are consumers).

The many roles that brands play in consumer decision making can materialize through multiple mechanisms, such as psychological (e.g., associative network memory), sociological (e.g., brand communities), and economic (e.g., brands as signals under uncertainty) processes (Keller 2002). Prior work on brands as signals under uncertainty (e.g., Erdem and Swait 1998; Wernerfelt 1988) focuses on the influence of imperfect and asymmetric information on consumer choice processes. Specifically, Erdem and Swait (1998) suggest that the clarity and credibility of brands as signals of product positions increase perceived quality, decrease consumer perceived risk and information costs, and thus increase consumer expected utility. They also discuss the antecedents to credibility and clarity. In this framework, the key characteristic of a brand signal is its credibility. "Brand credibility" (the credibility of a brand as a signal) is defined as the believability of the product position information contained in a brand, which depends on the willingness and ability of firms to deliver what they promise.

Most previous work using information economics frameworks to explain brand credibility and brand equity focuses only on the United States. The validity of existing frameworks in predicting consumer behavior in regard to brands as signals and outcomes outside the United States is not clear. With the growing trend in the globalization of marketing activities and the importance of brands in accomplishing many of these activities, there is a pressing need to address explicitly whether existing frameworks are compatible with consumer behavior outside the United States and, if not, to delimit their scope of application. Despite this practical impetus and growing theoretical interest, relatively little work examines the use of signals or extrinsic cues to judge quality across countries and/or cultures.

The purpose of this article is twofold. First, we test how well the information economics view of brand equity explains consumer brand choice and brand equity formation in different countries. Specifically, we assess the applicability of Erdem and Swait's (1998) framework across countries representing different cultural dimensions.[1] Second, we highlight some differences across countries. More specifically, we explore differences in the way brands as signals operate across countries and link any such differences to consumers' cultural orientation. We believe that three of Hofstede's (1980) cultural dimensions (collectivism/individualism, uncertainty avoidance, and power dis-

---

Tülin Erdem is E.T. Grether Professor of Business Administration and Marketing, Haas School of Business, University of California, Berkeley (e-mail: erdem@haas.berkeley.edu). Joffre Swait is a partner, Advanis Inc. (e-mail: Joffre_Swait@Advanis.ca). Ana Valenzuela is an assistant professor, Zicklin School of Business, Baruch College, City University of New York (e-mail: ana_valenzuela@baruch.cuny.edu). The authors thank the many colleagues who assisted with data collection: Esra Gençturk, Koç University in Istanbul, Turkey; Monica Gomez, Universidad Autonoma de Madrid; Nora Lado, Universidad Carlos III de Madrid; Shrilekha Parikh, College of Commerce and Economics, University of Mumbai, India; Oliver Heil, University of Mainz, Germany; Adele Diederich, Universität Oldenburg, Germany; Miyuri Shirai, Yokohama National University, Japan; André Menck, Universidade Federal de Uberlândia; and Cláudio Barbieri da Cunha, Escola Politécnica, Universidade de São Paulo, Brazil. This project was partially funded by a grant (SEJ 2004-00672/ECON) from the Spanish Ministry of Science and Education.

[1] Note that Erdem and Swait (1998) collected survey data on consumer self-reports of perceptions and purchase intentions and estimated a structural equation model (SEM) based on these survey data to assess the impact of brands as signals of expected utility, omitting price effects. In this article, using surveys, we collected consumers' self-reports of their perceptions and experimental data about their choices, and we estimated SEMs and brand choice models; for the latter, we also incorporated price effects.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 70 (January 2006), 34–49

tance) may particularly affect the way consumers use and think about brands as signals.

We conducted our analysis using survey and experimental data on orange juice and personal computers (PCs). We collected data on consumer brand perceptions, choices, and cultural dimensions from respondents in Brazil, Germany, India, Japan, Spain, Turkey, and the United States. We chose these countries to represent a wide range of Hofstede's (1980) cultural dimensions. All respondents were undergraduates in business schools (except for about half of the Brazilian sample, who were undergraduate engineering students) and participated in exchange for course credit. The results we obtained provide strong empirical evidence for brands as signals of product positions in general and for Erdem and Swait's (1998) framework in particular across countries that vary along the previously mentioned cultural dimensions. However, some interesting differences emerge in terms of both total credibility effects on choice and the mechanisms through which brand credibility effects on utility operate.

We organize the rest of the article as follows: In the next section, we review the relevant literature and discuss the conceptual framework. We then develop several hypotheses to be tested and describe the empirical approach and the data collection. The last section outlines the implications of our findings and discusses avenues for further research.

Literature Review and Conceptual Framework

Brands Under Uncertainty

The effect of uncertainty about brand attributes (in more general terms, a brand's position in the attribute space) on consumer choice has been the focal point of economics approaches to brand effects. Uncertainty affects consumers' perceptions of brand attributes, the variance of their attribute beliefs, and their information costs. The literature in both economics and marketing has focused on quality uncertainty, which is conceptualized and measured as a multidimensional construct (as a summary statistic of a brand's position in the space for attributes that are imperfectly observable). Aaker (1991) suggests that, all else being equal, strong brands are associated with higher perceived quality, which refers to mean beliefs about quality. Uncertainty about quality also implies that there is variance with consumer beliefs about quality for a given consumer (if consumers knew quality exactly or believed that they knew it exactly, the variance would be zero). This creates consumer perceived risk, and consumers tend to be risk averse in most contexts. Risk aversion affects consumers' decision making in various ways (e.g., Rao and Bergen 1992; Shimp and Bearden 1982). Risk-averse consumers feel threatened by ambiguous and uncertain product assessments. In this context, the literature suggests that brands can reduce perceived risk by becoming credible and consistent symbols of product quality (Erdem and Swait 1998; Montgomery and Wernerfelt 1992).

When there is quality uncertainty, especially in the presence of risk aversion, consumers tend to search for more information about product quality before making a decision (Money, Gilly, and Graham 1998; Shimp and Bearden 1982).[2] However, such information may often not be available or may not be credible if it is available. If it is available and credible, consumers may not be willing to bear large information costs, because utility decreases with increasing information costs. Consequently, consumers may resort to extrinsic cues, such as price, advertising, or channel choice, to infer product quality (Zeithaml 1988). For example, brand as a cue of quality could be particularly useful in countries (e.g., India) in which the quality of the products may vary widely in a given product category (Maxwell 2001).

The signaling literature in economics focuses on which conditions cause these cues to become credible signals (i.e., manipulatable attributes or activities of economic agents to convey information about their characteristics; see, e.g., Spence 1974). A firm can use various marketing-mix elements to signal product quality (e.g., charging a high price, offering a certain warranty, distributing through certain channels). Each of these actions may or may not be credible depending on market conditions, including competitive conditions and consumer behavior. However, what sets brands apart from the individual marketing-mix elements as credible signals is that the former embody the cumulative effect of prior marketing-mix strategies and activities. The historical notion that credibility is based on the sum of prior behaviors has been referred to as "reputation" in the information economics literature (see Herbig and Milewicz 1995).

Credibility is broadly defined as the believability of an entity's intentions at a particular time. Credibility is posited to have two main components: expertise and trustworthiness (Rao and Ruekkert 1994; Wernerfelt 1988). Thus, brand credibility is the believability of the product information contained in a brand, which requires that consumers perceive the brand as having the ability (i.e., expertise) and willingness (i.e., trustworthiness) to deliver continuously what has been promised (brands can function as signals because if and when they do not deliver what is promised, their brand equity will erode). All else being equal, the credibility of a brand has been shown to be greater for brands with greater marketing-mix consistency over time and greater brand investments (Erdem and Swait 1998). Consistency refers to the degree of harmony and convergence among the marketing-mix elements and the stability of marketing-mix strategies and attribute levels over time. Brand investments are resources that firms spend on brands to (1) assure consumers that brand promises will be kept and (2) demonstrate long-term commitment to brands (Klein and Leffler 1981). Furthermore, it has also been shown that the clarity (i.e., lack of ambiguity) of the product information contained in a brand is an antecedent to brand credibility (Erdem and Swait 1998).

As Aaker (1991) also suggests, higher perceived (or expected) quality, lower information costs, and lower risks associated with credible brands can increase consumer evaluations of brands. Using structural equation models (SEMs), Erdem and Swait (1998) show that expected utility increases with perceived quality and decreases with perceived risk and information costs; in turn, these are anteceded by brand credibility.

---

[2] Specifically, in cross-cultural contexts, consumers in cultures that rate high in collectivism and uncertainty avoidance have been shown to use personal information sources more often than those that rate low on these two cultural dimensions (Dawar, Parker, and Price 1996).

Brand Effects Across Countries and Cultural Dimensions

Few previous studies examine the use of signals or extrinsic cues to judge quality across countries and/or cultures. Dawar and Parker (1994) find that there are few differences in the use of price to signify quality for a highly homogeneous segment of consumers across both Western countries and Japan. McGowan and Sternquist (1998) compare Japanese and U.S. consumers in terms of their price–quality schema, prestige sensitivity, and value consciousness. The results indicate that price-inference behaviors may be market universal. However, these studies focus on comparing consumer price–quality perceptions across developed markets. In inefficient markets, such as those of less-developed countries, consumers may believe in the price–quality relationship to a lesser extent because price information is less credible (Zhou, Su, and Bao 2002). As a result, consumers in such markets may need to rely more on certain other signals to infer quality because the usual product information is less available or less reliable.

Prior research has also attempted to link brand choice to several cultural dimensions (e.g., Cray and Mallory 1998; Trompenaars and Hampden-Turner 1997; Voich 1995). Particularly, Hofstede's (1980, 1984) influential work on cross-cultural value systems identifies three aspects of cultures that can be related to brand choice: collectivism/individualism, uncertainty avoidance, and power distance.[3] Although Hofstede conducted his research in organizational settings, the values he identifies have often been associated with consumer behavior, such as consumer tipping behavior (Lynn, Zinkhan, and Harris 1993), country-of-origin effects on product evaluations (Gurhan-Canli and Maheswaran 2000), value–attitude relationships (Gregory, Munch, and Peterson 2002), persuasion effects (Aaker and Maheswaran 1997), consumer innovativeness (Steenkamp, Hofstede, and Wedel 1999), and behavioral intention models (Lee and Green 1991).

The marketing literature has been particularly keen in examining the cultural dimensions of collectivism and individualism. The collectivism/individualism dimension relates brand usage to social motivations (Hofstede 1980). Cultures that are high in individualism tend to seek variety and hedonistic experiences, whereas cultures that are high in collectivism correlate more with conformity and group behavior. Roth (1995) finds support for the hypotheses that collectivist cultures consider brands that reinforce group membership and affiliation more attractive, whereas individualistic cultures favor brands that reinforce their independence and provide individual gratification. Furthermore, collectivist societies value consensus, which makes them loyal to the dominant brand (Robinson 1996). Research into the symbolic meaning of brands indicates that brand names are an important symbol of group identity in collectivistic societies (Johansson, Ronkainen, and Czinkota 1994). Finally, previous literature has found that the degree to which other people influence the consumer purchase decision is a factor that may affect brand attitudes across national cultures. Particularly, Nicholls, Roslow, and Dublish (1997) show that a collectivist subculture (Hispanic consumers in the United States) tends to be more susceptible to social influence than an individualistic subculture (their Anglo counterparts). For example, collectivist consumers considered the influence of friends in the choice of a favored brand very important.

In this article, and given our focus on brands as signals, it is likely that the most relevant cultural mechanism is uncertainty avoidance. Uncertainty avoidance refers to "the extent to which people feel threatened by ambiguous situations and create beliefs and institutions that try to avoid these" (Hofstede and Bond 1984, p. 418). This concept captures the cultural pattern of seeking stability, predictability, and low risk rather than change and new experiences (Hofstede 1984). Consequently, consumers in cultures that are high in uncertainty avoidance tend to favor credible and consistent brands (Dawar and Parker 1994; Robinson 1996).

Another cultural dimension, power distance, describes the extent to which a culture fosters social inequality. Cultures that are high in power distance tend to emphasize prestige and wealth in shaping boundaries or vertical relationships between social and economic classes, such as rich/poor and superiors/subordinates (Hofstede 1984; Inkeles 1960). If the aspects of cultural power distance are related to consumer needs, it appears that social brand images should be the best fit for high-power-distance cultures. Prior research confirms that consumers in high-power-distance cultures attach more importance to products' brand names than consumers in low-power-distance cultures (Bristow and Asquith 1999; Robinson 1996; Roth 1995). However, when power distance is high, consumers also tend to "distrust" authority (Hofstede 1980). In turn, greater distrust of authority may diminish brand credibility effects, counterbalancing the stronger positive impact of the relationship between social status and brand names.

Applicability of the Information Economics Perspective to Brand Effects Across Countries and Cross-Cultural Differences in the Use of Brands as Signals

Erdem and Swait (1998) show that brand credibility (1) increases perceived quality, (2) decreases perceived risk, and (3) decreases information costs, all three of which (4) increase consumer (expected) utility. We expect these main relationships to hold across countries but to be moderated by cultural dimensions. Uncertainty avoidance is the cultural dimension that is most clearly related to brand effects as signals. Consumers from high-uncertainty-avoidance cultures are more likely to be risk averse and should have lower tolerance for ambiguity. This type of consumer tends to perceive more risk from product consumption and therefore adheres to historically tested patterns of behavior (Verhage, Yavas, and Green 1991). Thus, as we already indicated, consumers from high-uncertainty-avoidance cultures may use brands as signals more than consumers from low-uncertainty-avoidance cultures. In addition, perceived risk may be a relatively more important source of the brand signaling effect in such cultures. Specifically, we expect that the impact of brand credibility through perceived risk on consumer utility is greater in cultures with high uncertainty avoidance (Douglas and Craig 1997). However, we also expect that high uncertainty avoidance decreases sensitivity to information costs and increases willingness to collect information. Thus:

---

[3] Hofstede had a fourth dimension, masculinity, and later added time orientation as a fifth dimension. We chose not to focus on masculinity, because the scale he developed combines two separate (though related) components: (1) attitudes toward gender roles and (2) attitudes toward quality of life. No straightforward expectations can be generated about how these might affect brand credibility effects. We pretested the impact of time orientation. However, there was not enough variation across countries on this dimension to study it further.

H_1: Brand credibility's impact on consumer utility through perceived risk and information costs saved is larger than its impact through perceived quality under high uncertainty avoidance than under low uncertainty avoidance.

H_2: Brand credibility's overall impact on consumer brand choice increases with uncertainty avoidance.

The collectivist/individualist distinction is a more complex phenomenon in regard to its possible moderating effects on brand signals. Collectivist societies subordinate individual goals to the goals of a few large in-groups (Triandis and Gelfand 1998). In individualistic societies, the social fabric and group norms are much looser. Collectivism should result in more reliance on the brand choice of the group as a whole. Therefore, all else being equal, reliance on brands, especially in forming quality perceptions, may be more important in highly collectivist societies because the positive impact of credibility on perceived quality and the impact of perceived quality on choice could become elements that enhance belongingness to the group. However, culture conditions the way consumers search for product information (Dawar, Parker, and Price 1996). Collectivist cultures tend to exhibit more information search than individualist cultures, but they rely more on interpersonal information exchange or word of mouth (Money, Gilly, and Graham 1998), which reduces the impact of information costs. This may have two consequences: (1) a dampening of the total credibility effects and (2) brand credibility effects operating more strongly through perceived quality than through perceived risk and information costs.

We expect brand credibility's role in enhancing quality perceptions and, thus, belongingness to the group to dominate the dampening impact of reliance on interpersonal information exchange. We believe that this will be the net effect because the importance of belonging to a group is the driver of behavior in collectivist societies. Thus:

H_3: Brand credibility's impact on consumer utility through perceived quality is more pronounced than its impact through perceived risk and information costs in high-collectivist cultures than in low-collectivist cultures.

H_4: Brand credibility's overall impact on consumer brand choice increases with collectivism.

High-power-distance cultures tend to be hierarchic in their interpersonal relationships and decision making, whereas low-power-distance cultures tend to be egalitarian. Power distance influences active opinion leadership behaviors by causing consumers to distrust others (Dawar, Parker, and Price 1996). In addition, power distance should make consumers distrust brands. In this sense, enhancements on brand credibility and clarity could have a lower effect on perceived quality, perceived risk, and information costs saved. However, high-power-distance cultures tend to emphasize the importance of prestige and wealth in vertical relationships between social classes (Hofstede 1980). If this aspect of cultural power distance is related to consumer needs, it appears that brands as signals of positions that are related to social status should be most important in high-power-distance cultures because people are highly motivated by status and affiliation norms (Roth 1995). Thus, brands as signals may have a larger effect on perceived quality, perceived risk, and information costs saved in high-power-distance cultures than in low-power-distance cultures. Given the counteracting factors that determine the impact of power distance on brand credibility effects, power distance may not have an effect on how brand credibility affects choice, or if one of the two major mechanisms (distrust versus importance of authority) dominates, it may decrease or increase total brand credibility effects. Consequently, we do not propose any hypotheses about the directionality of the moderating effects of power distance and leave the open empirical question for further analysis.

Empirical Investigation of Brand Signaling Across Cultures

Using a sample of U.S. university students, Erdem and Swait (1998) establish the validity of brand signaling (specifically, through brand credibility) as the basis for positive brand effects on product evaluations. We argue that certain basic relationships they postulate would be moderated across different cultures, particularly in terms of three important cross-cultural constructs: uncertainty avoidance, collectivism/individualism, and power distance.

We collected comparable samples of respondents in seven different countries, which we specifically chose because they differ across these cultural constructs. Sekaran (1983) identifies two primary ways to achieve sample comparability: drawing nationally representative samples or selecting matched samples on the basis of some set of characteristics of interest. Because of budgetary constraints that prevented representative sampling, we matched samples by recruiting undergraduate business student respondents in all countries. (Because of recruitment difficulties in Brazil, we used engineering undergraduates for about half of the sample.) Although this is a potential limitation to the generalizability of our results, Erdem and Swait (1998) first tested their model using undergraduate student samples, and therefore we can best establish cross-cultural generalizability with samples that have similar characteristics. We also expect that differences in age, sociodemographics, relative income, and so forth, are small when we use undergraduate student samples, making possible a clearer attribution of substantive theoretical differences among countries to cultural differences. Finally, the use of student samples constitutes a conservative test of the impact of cultural constructs. University students are one of the most globally homogeneous segments of consumers, and thus, any cultural difference may be amplified if data are collected by means of a more inclusive group of respondents.

We made cross-country comparisons in two ways: First, we estimated and compared country-level SEMs; second, we used mixed multinomial logit (MNL) choice models based on an experimental design to test the role of brand and cultural constructs in the choice process. We selected two product classes for testing the robustness of brand signaling across cultures. We used orange juice to represent low-involvement, low-cost product categories, and we used PCs to represent high-involvement, high-price categories.

The Structural Model

Figure 1 presents our SEM. With the exception of the relative price (RP) construct, this is essentially the same model as that which Erdem and Swait (1998) implement. We postulate that brand credibility (Cr) increases perceived quality (PQ), which increases information costs saved (ICS), which decreases perceived risk (or, conversely, increases its complement, lower perceived risk [LPR]); in turn, we expect that these latter constructs increase product consideration and purchase (C&P).

In addition to Erdem and Swait's (1998) brand constructs, the model includes an RP construct. We included this for two reasons: First, Erdem and Swait do not include price in any form in the SEM, and we believed that its addition would be an interesting extension of their work to control for purely economic effects on C&P. Second, rather than price itself, which would be country and time specific, we opted to use RP to capture the relative price positioning of the brands in each country. The RP positioning is likely to be a function of brand credibility because consumers are likely to perceive the firm as charging a premium if their brand is strong and visible in the market. This leads to our expectation that the path from Cr to RP will be positive. However, we expect that the effect of RP on C&P is negative: The greater the relative price of the brand, the less the brand should be considered and, thus, the less attractive it should be.

The Measurement Model

We used a total of 19 items (see Table 1) to implement the measurement model that supports the SEM in Figure 1. We

FIGURE 1
The SEM Relating Brand Credibility to Product Utility

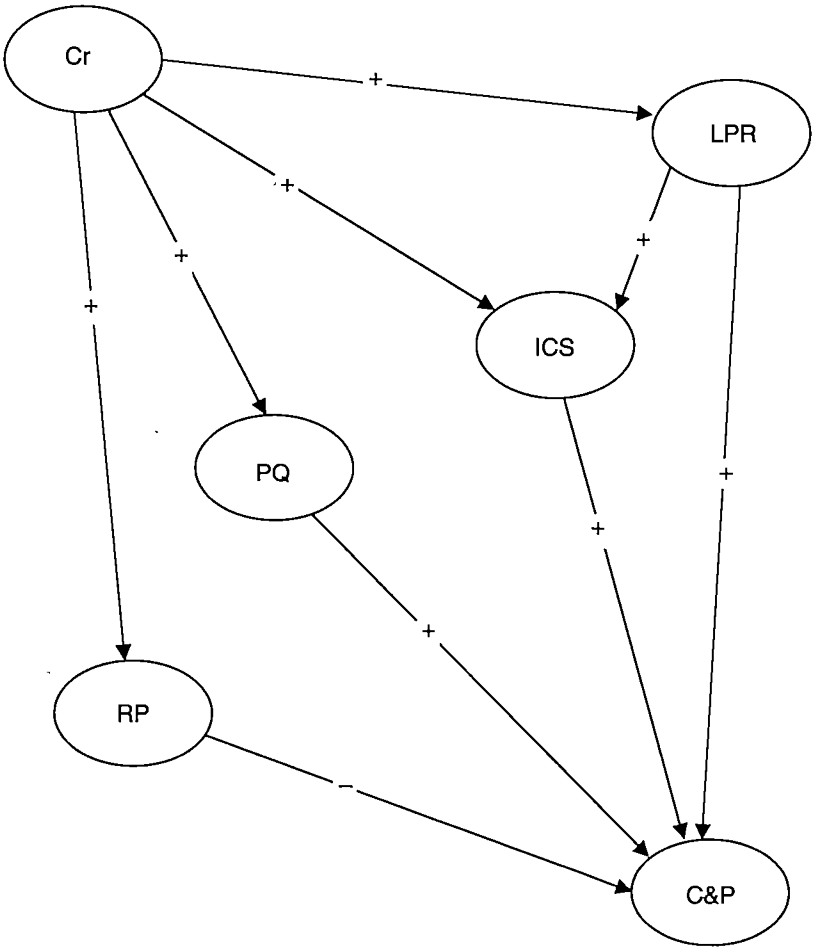


TABLE 1
Brand Constructs Measurement Model

Construct	Measurement Items	SMC Juice	SMC PC
Cr	1. This brand delivers what it promises.	.51	.56
	2. This brand's product claims are believable.	.41	.4
	3. Over time, my experiences with this brand have led me to expect it to keep its promises, no more and no less.	.41	.43
	4. This brand is committed to delivering on its claims, no more and no less.	.32	.36
	5. This brand has a name you can trust.	.53	.69
	6. This brand has the ability to deliver what it promises.	.54	.58
PQ	7. The quality of this brand is very high.	.72	.74
	8. In terms of overall quality, I'd rate this brand as a ...	.70	.73
ICS	9. Knowing what I'm going to get from this brand saves me time shopping around.	.30	.36
	10. This brand gives me what I want, which saves me time and effort trying to do better.	.79	.75
LPR	11. I need lots more information about this brand before I'd buy it. (R)	.39	.45
	12. To figure out what this brand is like, I'd have to try it several times. (R)	.45	.51
	13. I never know how good this brand will be before I buy it. (R)	.34	.51
RP	14. This brand is more expensive than the average brand in this category	.63	.69
	15. Compared to other brands, this brand is more expensive	.63	.70
C&P	16. I would never buy this brand. (R)	.30	.37
	17. I would seriously consider purchasing this brand.	.49	.53
	18. How likely would you be to purchase this brand?	.80	.74
	19. If you were to buy nine [product] at one time, how many of each brand would you buy to reflect your relative preference for each brand (assume these are the only brands available)?	.39	.33

Notes: We measured all items on nine-point "strongly disagree/strongly agree" scales, except Items 8 (nine-point "low quality/high quality" scale), 18 (nine-point "very unlikely/very likely" scale), and 19 (allocation of nine objects among brands to reflect preference). (R) after an item indicates that we reversed it for inclusion in the model. SMC = squared multiple correlation, or the percentage of item variance explained by the model; we based this number on the pooled sample across countries.

refer the reader to the work of Erdem and Swait (1998) for the items of all constructs except RP, which we specifically formulated for this study.

We established convergent and discriminant validity using Anderson and Gerbing's (1988) procedures. For convergent validity, pooling data across countries and product categories to yield a sample of 882 respondents and 5292 brand evaluations (as the procedure requires), we found that the construct coefficients in all item equations are statistically significant at the 95% level. This is reflected in Table 1 by the appreciable squared multiple correlation levels for all items. In terms of discriminant validity, we found that the hypothesis that the correlation between every pair of constructs is equal to one (i.e., there is no discrimination between the constructs in question) is rejected for all pairs of constructs in Figure 1 at a level of 95% or higher. The smallest chi-squared statistic is 34.1 (d.f. = 1), and it occurs between Cr and ICS, for which the estimated correlation is .91 (SE = .014). Thus, we established both convergent and discriminant validity at the pooled level.

Cultural Constructs

We used the 13 items in Table 2 to characterize the four cultural constructs of interest: uncertainty avoidance, collectivism/individualism, and power distance (Hofstede 1980). We developed these items using the following procedure: We used Hofstede's (1980) original organizational values survey items and Triandis's (1995) horizontal and vertical collectivism/individualism items, which we slightly modified to improve their applicability to a general consumer context. After we collected cultural identity data by country from our sample, we ran confirmatory factor analyses for each country to identify whether any item did not load high enough (our cutting point was factor loadings greater than .50) on a single factor. There were minimal discrepancies in the factor structure by country. Our final cultural construct specification excluded a total of 5 items from the original set (1 in the power-distance construct, 2 in the uncertainty-avoidance construct, and 2 in the collectivism/individualism construct) that did not load high overall on the specific cultural construct. This procedure ensured reasonable levels of scale reliability at both the overall level and country by country (all Cronbach's $\alpha > .60$). Table 2 shows the final factor loadings and reliability coefficients using data at the aggregate level. Henceforth, whenever we refer to these constructs in terms of inclusion in a statistical model, it should be understood that we operationalized the cultural constructs as the simple average of the respective item responses for each individual respondent.

Data Collection Methodology

The survey we administered to the respondents has several sections: (1) elicitation of the measurement model items (nine-point "strongly disagree/strongly agree" scale; see Table 1) for three brands in each of two product classes; (2) elicitation of the degree of confidence (seven-point "very low confidence/very high confidence" scale) in evaluating a new brand in each product class; (3) an association task linking ("yes/no" scale) product class to items describing familiarity, risks of purchasing, benefits offered, involvement levels, and so forth (Erdem, Swait, and Louviere 2002); (4) a brand/price experimental choice task for each

TABLE 2
Cultural Constructs

Construct	Items	Factor Loadings
UA (α = .692)	1. Security is an important concern in my life.	.772
	2. Life is so uncertain that one must continuously be on the alert so as not to be caught at a disadvantage.	.784
	3. It is important to consider dissenting views when making personal and social decisions.	.808
C/I (α = .735)	4. I like sharing little things with my neighbors.	.537
	5. Being a unique individual is important to me. (R)	–.804
	6. Decisions reached in groups are better than those reached by single individuals.	.587
	7. I usually sacrifice my self-interest for the benefit of my group.	.597
	8. I'd rather depend on myself than on others. (R)	–.728
	9. It is important to me to be useful to others.	.758
PD (α = .628)	10. One's boss is a person just like oneself. (R)	–.633
	11. Bosses are always inaccessible and distant.	.782
	12. The way to change a society is to make everyone equally powerful. (R)	–.619
	13. Other people are a threat to one's power and cannot be trusted.	.723

Notes: We measured all items on nine-point "strongly disagree/strongly agree" scales. (R) after an item indicates that we reversed it before use. We obtained the factor loadings from principal components single-factor analysis. PD = power distance, C/I = collectivism/individualism, and UA = uncertainty avoidance

TABLE 3
Country Ratings by Cultural Dimension (Hofstede 2003)

Country	PD	C/I[a]	UA
Brazil	69	38	76
Germany	35	67	65
India	77	48	40
Japan	54	46	92
Spain	57	51	86
Turkey	66	37	85
United States	40	91	46

[a] A higher score on the C/I dimension represents higher levels of individualism.

Notes: Hofstede analyzed a large database of IBM employees' value scores collected between 1967 and 1973 that covered more than 70 countries. In the editions of Hofstede's work since 2001, scores are listed for 74 countries and regions, partly based on replications and extensions of the IBM study on different international populations. To collect these scores, Hofstede used surveys that were designed for measuring culture-determined differences among matched samples of respondents from different countries and regions. It consists of 20 content questions and 6 demographic questions. The answers are grouped into cultural value scores Value scores range from 0 to 100. Scores closer to 100 denote stronger agreement with a particular cultural value. PD = power distance, C/I = collectivism/individualism, and UA = uncertainty avoidance.

product class, involving 17 choice sets that each elicited choice among three brands and a no-choice alternative (the experimental design used all brands present in every choice set, and the same four levels of price applied to all brands; this is the same design that Erdem, Swait, and Louviere [2002] use); (5) elicitation of the cultural constructs (nine-point "strongly disagree/strongly agree" scale; see Table 2); and (6) some simple demographics (age, sex, household income, citizenship status, and number of times traveled abroad).

Originally developed and tested in U.S. English, the final survey instrument was then translated and back-translated into Spanish, Portuguese, Japanese, Turkish, and German for respondents in Spain, Brazil, Japan, Turkey, and Germany to ensure comparability across countries (for a description of cross-country experimental controls, see the Appendix). Respondents in the United States and India were given the original U.S. English survey, duly customized for local conditions (e.g., brands, prices). In the case of India, our recruiter examined the survey for understandability. We selected these countries to represent a wide range of uncertainty-avoidance, collectivism/individualism, and power-distance dimensions (Hofstede 2003). Table 3 shows Hofstede's (2003) original country ratings by dimension for these countries.

In each country, we selected three brands in each product class (orange juice and PCs) to represent different levels of market power and price/quality positioning.[4] We required all brands selected to have high levels of consumer brand recognition. We also ensured that at least one local brand and one regional or global brand were always included. Thus, the brands used in the study could be local, regional, national, or multinational/global, but to the extent possible, we selected them to cover the spectrum of market share (lower, average, higher), thus introducing greater potential heterogeneity in the brand constructs of interest. In the choice experiments, we used country- and product-specific price levels expressed in the local currency. We established average price levels across stores using price audits by country with the cooperation of our recruiters.

We then programmed the surveys for execution on the Internet, which ensured method invariance. We provided recruiters at each institution of higher learning with a list of identifiers and corresponding passwords (to enhance survey security), which they provided to their students for voluntary participation in the survey. In general, these recruiters were instructors of undergraduate business courses, except in Brazil, where we used both business and engineering undergraduates because of recruitment difficulties. The respondents have similar demographic profiles: They are young (average age varies from 18 years in India to 25 years in Germany; 79% of all respondents are 22 years of age or younger) and mainly middle class. The gender split varies somewhat more than age: The male/female split is about even in the United States and India; there are approximately 40% males/60% females in Germany, Spain, and Turkey; and there are approximately 55% males/45% females in Brazil and Japan. We excluded foreign exchange students from the sample. Students were given course credit for survey completion and were urged to complete the survey in one sitting (approximately 30 minutes); however, if the session was interrupted, respondents could subsequently resume the survey at the point of interruption.

---

[4] We also controlled for the potential of certain brands to communicate more individualistic values and others to communicate social values. In addition, we believe that such discrepancy is not as prevalent in the two product categories under study.

Data collection occurred from October 2002 to April 2003, depending on the country. This extended data collection period was necessary to obtain reasonable sample sizes for our statistical tests. Despite our best efforts, sample sizes by country vary significantly because of factors outside our control: Brazil (126 respondents), Germany (57), India (161), Japan (114), Spain (155), Turkey (126), and the United States (143).

Country-Specific SEMs

Table 4 presents the standardized SEM coefficients for each country and product class and the averages of cultural constructs for each country. We used the Amos 5 software program to obtain model estimates (Byrne 2001). Overall, the models display a modest, though acceptable, level of goodness-of-fit (for details, see Table 4). All statistically significant coefficients in both product classes are in the expected direction, with a single exception: The LPR → ICS path, which we expected to have a positive coefficient, has a negative and statistically significant effect in Japan for PCs. Otherwise, it seems that, in general, the model structure from Figure 1 holds across the different countries.

That is not to say that the countries are homogeneous in terms of the path coefficients. In the juice category, for example, the impact of Cr on C&P is through the PQ construct in Brazil and Turkey. In Spain, however, the ICS and LPR constructs are also mediators of the impact of Cr on C&P, indicating that Spanish respondents displayed a more nuanced response to the brand in the evaluation of product utility, considering the role of the brand in risk reduction and information acquisition enhancements to the utility of the brand. India, the United States, and, to a lesser extent, Germany are more similar to Spain than to Brazil and Turkey in overall structure in the juice category. Note that, in general, the RP construct is not statistically significant at the 95% significance level in this product category, except for in Spain. In the PC category, PQ is again the main mediator between Cr and C&P for Brazil, Germany, and Turkey. The impact of Cr on C&P in India, Japan, Spain, and the United States occurs through paths involving LPR and ICS.

The total impact of Cr on C&P, as estimated through the SEMs in Table 4, is remarkably consistent across countries: Brazil (.806, .738 for juice and PC, respectively), Germany (.815, .732), India (.847, .831), Japan (.773, .66), Spain (.724, .712), Turkey (.862, .880), and the United States (.803, .855). The smallest impact for juice is found in Spain, and for PCs, it is found in Japan. In both categories, Turkey displays the largest impact of Cr on C&P (and India is a close second in both categories). As we pointed out previously, however, these effects and their magnitudes come about through different pathways in Figure 1.

To undertake a more systematic exploration of between-country differences in the path coefficients, we plotted the estimated coefficients as a function of cultural identity constructs. The averages of cultural constructs (see bottom of Table 4) suggest that Germany (M = 5.31) scores lowest on uncertainty avoidance, whereas Brazil (M = 6.72) and Spain (M = 6.57) score highest. Among the seven countries we studied, Spain (M = 5.18) and Brazil (M = 5.22) scored highest on collectivism, whereas India (M = 4.55) and the United States (M = 4.69) scored lowest. Power distance seems to be lowest in Germany (M = 3.88) and highest in Japan (M = 4.99). These ratings are consistent with Hofstede's (2003) original findings (see Table 3), except for India and Turkey, which seem to rate more moderately in our current data than in Hofstede's work. A possible reason for this divergence is that we collected these cultural values using university students, who are more likely to have an international mind-set, and therefore ratings may tend to equalize those of reference countries in the Western world, such as the United States and Europe.

Figure 2 shows three plots that depict the coefficient estimates for the Cr → LPR → ICS → C&P paths as a function of uncertainty avoidance (UA) averages for each country. These paths are shown because they depict UA's moderating pattern relatively clearly: The higher the country on the UA scale, the stronger the first two paths become. This implies that as UA increases, Cr lowers risk perceptions more strongly and leads to greater ICS. This impact on risk is mitigated in part by the decreased reliance on the brand to save on information costs in decision making, as is evidenced by the downward trend in the strength of the ICS → C&P path as a function of increased UA (because UA decreases sensitivity to information costs). Furthermore, the PQ → C&P path strength does not vary systematically with UA. These combined results suggest that the moderating role of UA in regard to the impact of Cr on C&P is more through the LPR → ICS path than through the PQ path. Thus, we find evidence in support of H_1.

Figure 3 shows two plots of the coefficient estimates for the Cr → PQ → C&P paths as a function of the collectivism/individualism (C/I) construct. These paths depict a different moderating pattern for C/I. Although Cr → PQ path coefficients do not change significantly, there is a discernible positive trend in PQ → C&P path coefficients with increasing collectivism (Spain is an exception to this pattern). Conversely, the sequence of paths Cr → LPR → ICS → C&P is also found to be a function of the C/I construct. The higher a country scores on the C/I scale (implying a more collectivist culture), the larger are the standard-

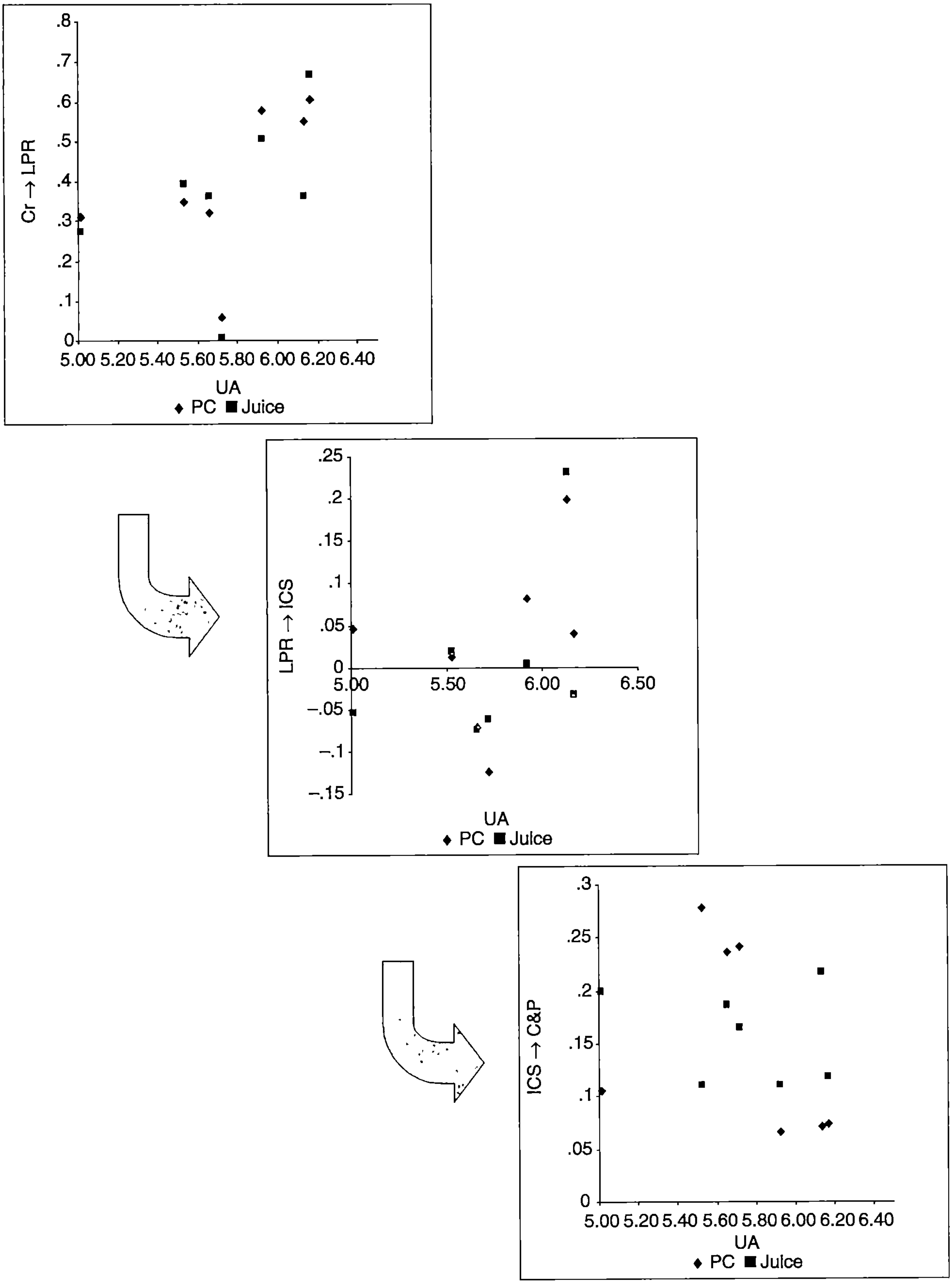
TABLE 4
SEM Estimation Results by Country

Path	Standardized Coefficients						
	Brazil	Germany	India	Japan	Spain	Turkey	United States
PCs							
Cr →LPR	.612**	.314**	.325**	.057	.559**	.579**	.347**
Cr →PQ	.851**	.885**	.953**	.798**	.849**	.973**	.972**
Cr →ICS	.801**	.564**	.848**	.69**	.728**	.703**	.735**
Cr →RP	.538**	.488**	.768**	.538**	.815**	.754**	.455**
LPR →ICS	.034	.053	–.074	–.127*	.189**	.1	.014
LPR →C&P	.033	.023	.194**	.102*	.144**	.01	.01
ICS →C&P	.137	.117	.231**	.242**	.134	.051	.216**
RP →C&P	–.18**	–.109	.004	–.188**	–.096	.068	–.043
PQ →C&P	.825**	.803**	.603**	.74**	.704**	.801**	.732**
Juice							
Cr →LPR	.664**	.276**	.365**	.014	.375**	.507**	.382**
Cr →PQ	.968**	.879**	.95**	.937**	.877**	.972**	.955**
Cr →ICS	.76**	.775**	.827**	.67**	.7**	.805**	.674**
Cr →RP	.644**	.489**	.416**	.514**	.454**	.76**	.604**
LPR →ICS	–.022	.031	–.076	–.072	.219**	–.016	.03
LPR →C&P	.004	–.013	.116**	–.005	.092	.082	.129**
ICS →C&P	.098	.262**	.178**	.134	.231**	.09	.105
RP →C&P	–.082	.059	.017	–.104	–.124**	–.083	–.052
PQ →C&P	.808**	.664**	.69**	.786**	.645**	.836**	.747**
Number of respondents	126	57	161	114	155	126	143
Number of brand observations	378	171	483	342	465	378	429
Cultural Constructs (Averages)							
UA	6.16	5.01	5.66	5.72	6.13	5.92	5.53
C/I[a]	5.41	5.03	4.90	5.07	5.54	5.29	5.05
PD	4.03	3.82	4.16	4.82	4.01	4.27	4.51

*$p < .05$.
**$p < .01$.
[a] A higher score on the C/I dimension represents higher levels of collectivism
Notes. χ^2 =11333.6, d.f. = 2052; χ^2/d.f. = 5 52, goodness-of-fit index = 83, comparative fit index = .83, root mean residual = .470, and root mean square error of approximation = .029. PD = power distance, C/I = collectivism/individualism, and UA = uncertainty avoidance.

FIGURE 2

Plot of Cr → LPR → ICS → C&P SEM Coefficients as a Function of UA

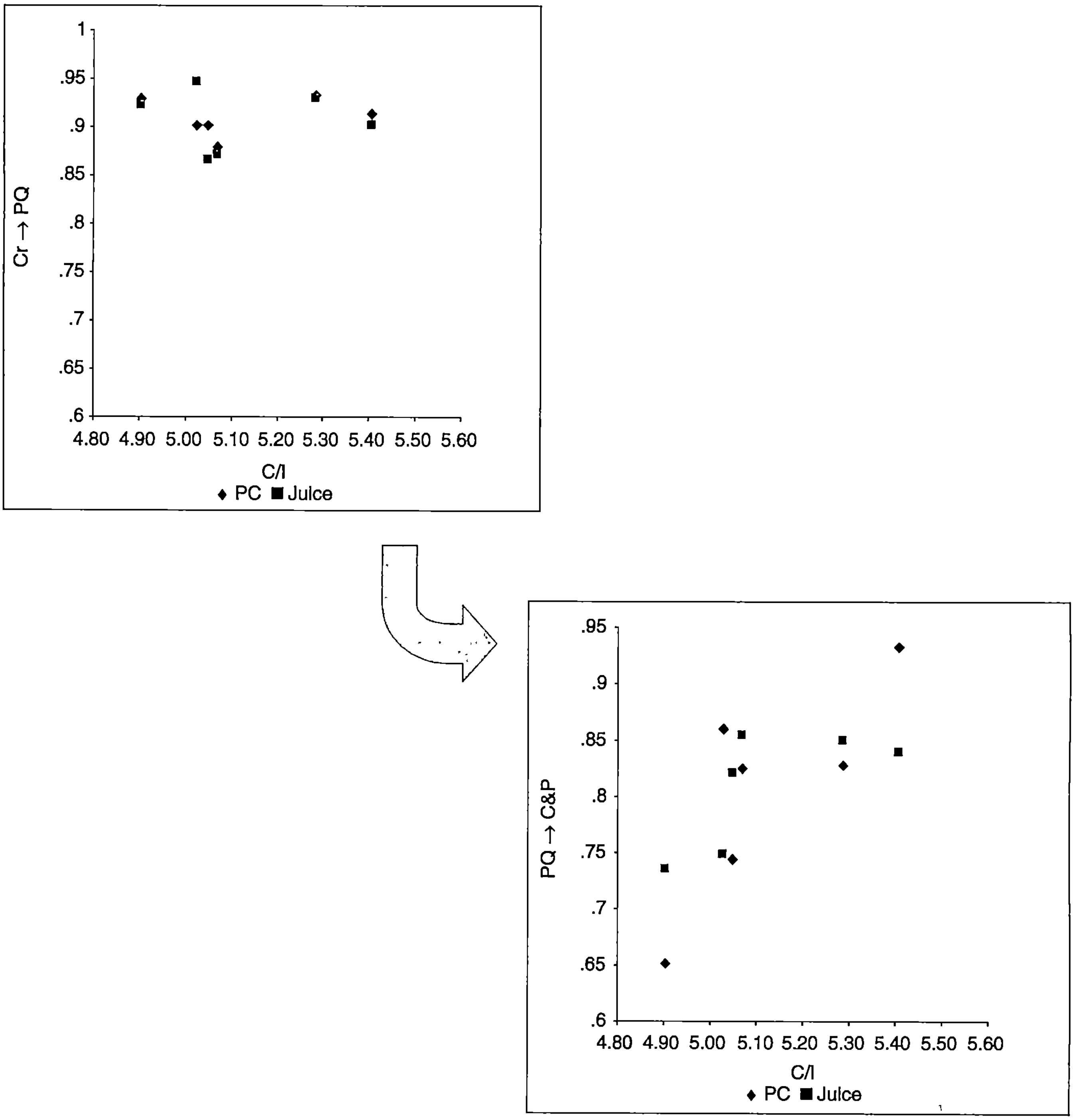


ized coefficients for Cr → LPR and LPR → ICS, and the smaller is the ICS → C&P path coefficient, with an almost null overall effect. These results suggest that in collectivist countries, the impact of Cr is felt more directly through the PQ route than through the information acquisition route (LPR → ICS). This seems intuitive because it is reasonable that in collectivist cultures, sources of information, such as word of mouth, may be more easily available and more often used. Thus, we find evidence in support of H_3. Notably, the last cultural construct under study (power distance [PD]) does not seem to explain any of the differences between path coefficients.

Evidence from the Choice Modeling Results

To analyze the overall impact of credibility on choice in various countries (while controlling for price effects) and to

FIGURE 3
Plot of Cr → PQ → C&P SEM Coefficients as a Function of C/I (Except Spain)



test H_2 and H_4, we also estimated brand choice models. The final statistical results we present are based on mixed MNL (McFadden and Train 2000) specifications, using the experimental choice data we mentioned previously. These data resulted from eliciting 17 choices for each product class. We always presented three brands in each country and product category and a no-choice alternative; each brand was associated with a four-level price variable, which had the same range for all brands in the same country and product class. Note that 17 choice sets are more than what is considered minimally necessary for identification of parameters, but we did this to maintain compatibility and comparability with prior studies (especially Erdem, Swait, and Louviere 2002).

In this section, we examine whether the impact of credibility on preference differs in magnitude between countries or, more specifically, whether the impact of brand credibility on choice is moderated by the cultural constructs used in the previous two sections. To this end, we calibrate two mixed MNL models, one per product category, which allow for stochastic heterogeneity of brand and price effects at the country level. The utility function also includes the Cr construct and its interactions with UA, C/I, and PD. Precisely, the utility functions are given by the following expression:[5]

$$(1)\ U_{icn} = \alpha_{icn} + \beta_{icn} \times LN(p_{ic}) + \delta \times Cr_{icn} + \gamma_1 \times UA_{icn} \times Cr_{icn} + \gamma_2 \times (C/I)_{icn} \times Cr_{icn} + \gamma_3 \times PD_{icn} \times Cr_{icn} + \varepsilon_{icn},$$

where i, c, and n are index brand, country, and respondent, respectively; p_{ic} is a country-specific price for the good in question; ε_{icn} is the brand–country–person stochastic utility, which we assume to be independent and identically Gumbel distributed across brands and respondents within country but is allowed to have different scale/variance by country (Swait and Louviere 1993); α, β, δ, and γ are parameter vectors; and all other quantities have been previously defined. Note that the α and β parameter vectors are subscripted by respondent, reflecting that these coefficients are allowed to have stochastic distributions (we provide more details with results in Table 5).

We did not give the coefficient vectors for Cr and its cultural construct interactions stochastic distributions, and we assumed that they were comparable across countries. To support this specification, we established that the Cr construct has partial metric invariance across countries at the 95% confidence level by using Steenkamp and Baumgartner's (1998) test; this result holds for both product categories. The Cr construct has six measurement items in total; we established partial metric invariance with, at most, two items being country specific. This test shows that the construct is homogeneous and robust across the countries tested (for orange juice: $\chi^2 = 26.0$, d.f. = 18, $p = .10$; for PCs: $\chi^2 = 34.4$, d.f. = 23, $p = .06$). This also implies that comparing credibility scores across countries is empirically supported in our data, so making the γ and δ parameters in Equation 1 equal across countries is consistent with empirical evidence.

The resulting product class models appear in Table 5. Note that there is strong support for Erdem and Swait's (1998) contention that Cr increases product attractiveness/utility: Across all countries, there is a positive and significant main effect of Cr in both product classes. This result is robust because we allow for substantial country- and respondent-level differences through brands, prices, and stochastic taste distributions at the country level. Note that because of scale differences (i.e., error variance differences) across product classes, it is not appropriate to compare juice and PC coefficients directly; only parameter ratios or elasticities should be compared. The coefficients of substantive interest here are the interactions between Cr and the cultural constructs (i.e., δ, γ_1, ..., γ_4). In both product classes, we found that PD has no substantive impact on the main effect of Cr on utility, because γ_3 and γ_4 are not significantly different from zero at the 95% significance level; this is in agreement with prior analyses we previously reported, but it is based on SEMs. Of greater interest is the consistent finding across both product classes that UA increases the impact of Cr in both the juice and the PC categories (coefficient γ_1 in both categories), as we previously argued. The marginal impact of UA is relatively greater for PCs than for juice, as would seem reasonable, but it is statistically significant and increasing in both. Thus, we found evidence in support of H_2 in both the PC and the juice categories. Finally, increasing C/I also leads to the increased importance of Cr but only in the juice category (see coefficients γ_2 in Table 5). A possible reason for this finding is that brands as symbols of collective behavior are more often embraced in high-frequency, low-involvement purchases for which decision time pressure is high and the financial risk of making a wrong choice is low. Thus, we find evidence in support of H_4 only in the juice category. Note that the impact of C/I is significantly greater than that of UA in the juice category. It is difficult to discern the relative importance of these two cultural constructs across product classes because they are subject to different error term scales.

Discussion, Conclusions, and Further Research

In this article, we investigated the effects of brand credibility on consumer choice through perceived quality, perceived risk, and information costs saved across seven countries in Asia, Europe, and North and South America. These countries have been shown to vary across the cultural dimensions of uncertainty avoidance, collectivism, and power distance (Hofstede 1980). We found strong support for Erdem and Swait's (1998) framework of brand effects on choice under consumer uncertainty. More specifically, we found that brand credibility, the key construct in information economics–based approaches to brand equity, affected consumer choice in all the countries we studied. Although the paths through which credibility operated differed somewhat across countries, we found that the general framework of

---

[5] Note that an alternative utility specification would include the successors of credibility only (i.e., perceived quality, perceived risk, and information costs saved). Previous research has shown that credibility does not have a direct effect on utility, and its effects are mediated through these three antecedents (Erdem and Swait 2004). Therefore, there is no need to put credibility and the antecedents together in the utility function. We decided to put credibility rather than its antecedents in the utility to assess the total impact of Cr on choice and obtain a clearer and simpler picture of total Cr effects across countries.

TABLE 5
Mixed MNL Parameter Estimates

Utility Function[a]	Juice (Asymptotic t-Statistics)[b]	PCs (Asymptotic t-Statistics)[c]
Brazil		
Brand 1	−2.73 (−14.9)|3.643 (6.96)	44.569 (12.66)|.201 (5.47)
Brand 2	−.571 (−2.78)|.789 (5.26)	44.223 (12.67)|.201 (—)
Brand 3	−3.241 (−16.61)|4.727 (6.42)	43.246 (12.67)|1.718 (6.03)
−LN(price)	1.784 (27.03)|.371 (10.73)	—
LN(price)	—	−5.875 (−12.87)|.049 (5.15)
Germany		
Brand 1	−36.172 (−1.23)|1739.369 (.56)	44.449 (57.78)|4.6 (6.27)
Brand 2	3.723 (.63)|234.583 (.55)	46.086 (58.38)|2.706 (5.66)
Brand 3	−46.155 (−1.22)|641.065 (.55)	43.053 (57.02)|4.853 (3.72)
−LN(price)	3.815 (4.23)|.996 (6.66)	—
LN(price)	—	−5.952 (−60.46)|.053 (2.41)
India		
Brand 1	4.881 (13.28)|1.155 (6.18)	29.546 (10.78)|3.007 (5.03)
Brand 2	6.106 (13.83)|1.673 (9.06)	31.852 (10.78)|.391 (5.21)
Brand 3	6.189 (13.61)|.117 (3.83)	31.211 (10.77)|.391 (—)
−LN(price)	1.188 (21.93) |.067 (15.27)	—
LN(price)	—	−3.586 (−11.21)|.09 (4.95)
Japan		
Brand 1	1.188 (21.93)|.067 (15.27)	22.384 (12.85)|.122 (4.97)
Brand 2	26.54 (17.74)|.067 (—)	22.036 (12.84)|.432 (5.52)
Brand 3	26 066 (17.73)|1.766 (8.42)	22.22 (12.87)|.234 (5.13)
−LN(price)	1.881 (36.1)|.004 (12.84)	—
LN(price)	—	−3.439 (−13.25)|.037 (5.82)
Spain		
Brand 1	7.482 (14.47)|.356 (6.87)	8.605 (6.96)|.118 (2.42)
Brand 2	6.863 (14.01)|2.934 (7.17)	7.444 (6.27)|2.125 (4.18)
Brand 3	7.974 (15.12)|.356 (—)	8.736 (6.96)|.023 (1.01)
−LN(price)	.886 (16.55)|.049 (11.03)	—
LN(price)	—	−1.198 (−7.85)|.056 (2.88)
Turkey		
Brand 1	64.338 (31.79)|.204 (2.91)	22.577 (56.94)|.095 (3.74)
Brand 2	65.109 (31.79)|.886 (11.48)	24.321 (54.12)|.095 (—)
Brand 3	64.384 (31.6)|.114 (2.97)	24.499 (53.09)|.36 (6.67)
−LN(price)	1.484 (54.48)|.262 (8.36)	—
LN(price)	—	−2.581 (−37.19)|1.398 (13.06)
United States		
Brand 1	−.664 (−2.56)|1.607 (5.84)	13.147 (10.79)|.467 (3.82)
Brand 2	−1.188 (−5.47)|1.055 (4.95)	13.374 (10.73)|.374 (4.82)
Brand 3	−.398 (−1.47)|.409 (3.86)	13.318 (10.65)|.149 (3.09)
−LN(price)	1.716 (23.2)|.529 (10.46)	—
LN(price)	—	−2.148 (−11.92)|.045 (4.26)
Cr (δ)	.438 (6.47)	.448 (5.4)
Cr × UA (γ_1)	.009 (2.03)	.025 (5.28)
Cr × C/I (γ_2)	.064 (7.84)	.000 (.03)
Cr × PD (γ_3)	−.002 (−.41)	−.006 (−1.02)
LN(Scale Function)		
Brazil	−.193 (−3.12)	.581 (7.31)
Germany	−1.498 (−1.68)	.0 (—)
India	−.125 (−2.49)	−.115 (−1.3)
Japan	.238 (4.45)	.661 (8.7)
Spain	.087 (1.66)	.164 (1.71)
Turkey	.0 (—)	.0 (—)
United States	.862 (12.32)	.813 (9.76)

TABLE 5
Continued

Utility Function[a]	Juice (Asymptotic t-Statistics)[b]	PCs (Asymptotic t-Statistics)[c]
Goodness-of-Fit		
Log-likelihood at 0	–20,786.1	–20,758.4
Log-likelihood at convergence	–10,339.1	–11,985.0
Rho-squared	.5026	.4226
Number of respondents	882	881
Number of choices	14,994	14,974

[a] Prices are in currency of the respective country

[b] Heterogeneity specification: Numbers represent independent normal distributions for brand effects and independent log normal distributions for prices. For brand effects, presentation format is mean|variance of distributions, and t-statistics are in parentheses; for price effects, presentation format is mean|variance of the normal distribution of the natural logarithm of taste coefficients, and t-statistics are in parentheses To use the latter coefficients in the utility function, note that the taste coefficients will multiply the *negative* of price. We used the log normal distribution for this category's price coefficients because the specification led to higher goodness-of-fit than the normal distribution.

[c] Heterogeneity specification: Numbers represent independent normal distributions for brand and price effects. Presentation format is mean|variance of distributions, and t-statistics are in parentheses.

Notes: We used 100 Halton quasi-random numbers for simulation estimator.

brand credibility effects was empirically robust across the countries we studied.

To explore whether cultural constructs that have been linked to consumer choice behavior and brand effects in previous research moderated these effects, we investigated whether uncertainty avoidance, collectivism, and power distance explain any differences across the countries in regard to mechanisms through which brand credibility effects on consumer brand choice materialize. We expected that uncertainty avoidance and collectivism would increase total credibility impacts, and our results confirmed this. Given counterbalancing considerations in the case of power distance, we deemed its net effect to be an empirical question; we did not find any moderating effects of this construct in either analysis (SEM or choice model). Note, however, that this does not mean that this construct does not influence brand credibility effects. It merely indicates that given that many factors are at play in confounding directions, these factors may cancel out, and it may not be possible to observe any total net effect. Indeed, a necessary area for further research is to map out these different factors more carefully by measuring and testing them separately, for example, with the inclusion of a new construct that denotes brand perceived prestige (Steenkamp, Batra, and Alden 2003).

The results we obtained through the choice models estimated suggest that uncertainty avoidance amplifies the net effect of credibility on choice. We found that the effect was stronger for PCs than for juice, which we expected because PC is a more complex, higher-involvement product category than juice. We found that collectivism amplified the brand credibility effects as well, but only in juice. This might be due to the conspicuous consumption nature of juice, a more frequently purchased, low-ticket item for which a brand's value as a group identity symbol would be more important than for PCs, for which functional features prevail.

The results we obtained through the SEMs suggest that as uncertainty avoidance increases, brand credibility lowers risk perceptions and increases information costs saved. However, it also decreases the reliance on the brand to save on information costs. The same pattern of influence is true as collectivist tendencies increase. We believe that uncertainty avoidance and collectivism decrease the impact of information costs on consumer choice because of the increased willingness of consumers to engage in active search (in case of uncertainty avoidance) and the ability to use other low-cost information sources, such as word of mouth (in the case of collectivism). Conversely, increasing collectivism leads to a stronger path for the effect of perceived quality on brand consideration and purchase. This is consistent with the notion that collectivist consumers perceive credible brands as better quality and prefer them because they reinforce their belongingness to the group. Overall, these findings suggest that collectivism moderates the impact of brand credibility on consumer utility (consideration and purchase) by increasing the effect of perceived quality, whereas uncertainty avoidance causes the credibility impact to operate mainly through the reduction of perceived risk and information costs.

These findings have important managerial implications. Returns on a company's efforts to establish brand credibility depend on consumers' cultural values. Establishing a brand's credibility seems to be especially effective when consumers have more aversion to risky choices and, thus, consider brands a way to choose "the sure thing" and when consumers strongly need to reinforce their belongingness to a group through the consumption of symbols that reflect such identity. In addition, companies could take advantage of cultural differences in brand equity formation in many ways. For example, companies could work with product positionings that match cultural values (e.g., emphasizing brand attributes that invoke either risk reduction or social acceptance). Companies could also execute communication campaigns that reinforce a brand's ability to either reduce risk or generate group identification (e.g., showing brand consumption occasions in which such ability comes into play). These communication campaigns could also add elements that represent a brand's trustworthiness (e.g., messages with ideas such as "you can always count on brand X" for high-uncertainty-avoidance countries or "you will always be part of it" for high-collectivist countries). Finally, brand extensions could be perceived as categories in which either risk or collective identity matters.

There are many avenues of further research to pursue. Because the original body of empirical research in the United States is based on student samples and because taking relatively homogeneous samples across countries makes it easier to control for factors such as age, income, education, and other items that may lead to between-country differences, our study constitutes a conservative analysis of cultural differences. We expect that in more heterogeneous consumer samples, the effects of cultural constructs would be amplified. Testing the model on more heterogeneous consumer samples and with other product categories will also enhance the generalizability of our results. Another avenue for further research would be to investigate the influence of cultural dimensions other than Hofstede's (1980), which have been criticized on the basis of their descriptive nature and organizational focus, in the validity and applicability of the model.

In this study, we focused on brand credibility and its antecedents. However, factors that underlie brand credibility (e.g., temporal consistency of brand positioning) and subdimensions of such factors may be moderated by cultural constructs as well. A better understanding of underlying dimensions of consistency, brand investment perceptions, and the like has important branding and brand management implications in international markets. Indeed, the success of globalization strategies depends on assumptions about the communalities in mechanisms by which brands affect consumers, regardless of their cultural identity.

APPENDIX
Cross-Country Experimental Controls

Functional equivalence (Sekaran 1983)	To control for functional equivalence, our survey requested responses to similar problems in every culture. There was a comparable context or work setting in every country.
Survey instrumentalization (Steenkamp, Hofstede, and Wedel 1999)	We used back-translation to control for vocabulary equivalence, idiomatic equivalence, and grammatical and syntactical equivalence in our survey.
Conceptual equivalence (Adler 1983)	To control for conceptual equivalence, we pretested equivalence of the brands selected for the analysis and chose product categories that are purchased often, are equally available, and have the same function for consumers in each country of interest.
Construct equivalence (Steenkamp and Baumgartner 1998)	Although we do not present partial metric invariance because of space limitations, we established it using multiple-group SEMs for a subset of countries. This suggests cross-country comparability of path coefficients.
Response equivalence (Adler 1983)	We controlled for response equivalence by using uniform data collection procedures (Internet collection) and identical instructions.
Subject pool equivalence (Alden, Steenkamp, and Batra 1991)	We controlled for the equivalence in educational background in three ways: First, the universities we chose were all top-tier universities in their countries. Second, participants were all students in business or engineering. Third, we analyzed sample characteristics to verify the match in terms of age, sex, and social class.
Currency equivalence (Roth et al. 1991)	We controlled for currency equivalence by using organic prices in each country (i.e., current price averages across retail stores).

REFERENCES

Aaker, David A. (1991), *Managing Brand Equity*. New York: The Free Press.

Aaker, Jennifer L. and Durairaj Maheswaran (1997), "The Effect of Cultural Orientation on Persuasion," *Journal of Consumer Research*, 24 (4), 315–28.

Adler, N.J. (1983), "Cross-Cultural Management Research: The Ostrich and the Trend," *Academy of Management Review*, 8 (2), 226–32.

Alden, Dana L., Jan-Benedict E.M. Steenkamp, and Rajeev Batra (1999), "Brand Positioning Through Advertising in Asia, North America, and Europe: The Role of Global Consumer Culture," *Journal of Marketing*, 63 (January), 75–87.

Anderson, J. and D. Gerbing (1988), "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin*, 103 (3), 411–23.

Bristow, Dennis N. and Jo Ann L. Asquith (1999), "What's in a Name? An Intra-Cultural Investigation of Hispanic and Anglo Consumer Preferences and the Importance of Brand Name," *Journal of Product and Brand Management*, 8 (3), 185–97.

Byrne, Barbara M. (2001), *Structural Equation Modeling with Amos: Basic Concepts, Applications, and Programming*, 2d ed. Mahwah, NJ: Lawrence Erlbaum Associates.

Cray, David and Geoffrey R. Mallory (1998), *Making Sense of Managing Culture*. London: International Thomson Business Press.

Dawar, Niraj and Philip Parker (1994), "Marketing Universals: Consumers' Use of Brand Name, Price, Physical Appearance, and Retailer Reputation as Signals of Product Quality," *Journal of Marketing*, 58 (April), 81–96.

———, ———, and Lydia M. Price, (1996), "A Cross-Cultural Study of Interpersonal Information Exchange," *Journal of International Business Studies*, 27 (3), 497–517.

Douglas, S.P. and C.S. Craig (1997), "The Changing Dynamic of Consumer Behavior: Implications for Cross-Cultural

Research," *International Journal of Research in Marketing*, 14 (4), 379–95.

Erdem, Tülin and Joffre Swait (1998), "Brand Equity as a Signaling Phenomenon," *Journal of Consumer Psychology*, 7 (April), 131–57.

——— and ——— (2004), "Brand Credibility and Its Role in Brand Choice and Consideration," *Journal of Consumer Research*, 31 (1), 191–99.

———, ———, and Jordan Louviere (2002), "The Impact of Brand Credibility on Consumer Price Sensitivity," *International Journal of Research in Marketing*, 19 (1), 1–19.

Gregory, Gary D., James M. Munch, and Mark Peterson (2002), "Attitude Functions in Consumer Research: Comparing Value-Attitude Relations in Individualist and Collectivist Cultures," *Journal of Business Research*, 55 (11), 933–45.

Gurhan-Canli, Zeynep and Durairaj Maheswaran (2000), "Cultural Variations in Country of Origin Effects," *Journal of Marketing Research*, 37 (August), 309–317.

Herbig, Paul and John Milewicz (1995), "The Relationship of Reputation and Credibility to Brand Success," *Journal of Consumer Marketing*, 14 (4), 5–10.

Hofstede, Geert (1980), *Culture's Consequences: International Differences in Work-Related Values*. Beverly Hills, CA: Sage Publications.

——— (1984), "The Cultural Relativity of the Quality of Life Concept," *The Academy of Management Review*, 9 (3), 389–99.

——— (2003), *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*, 2d ed. Thousand Oaks, CA: Sage Publications.

——— and Michael Harris Bond (1984), "Hofstede's Cultural Dimensions: An Independent Validation Using Rokeach's Value Survey," *Journal of Cross Cultural Psychology*, 15 (4), 417–33.

Inkeles, Alex (1960), "Industrial Man: The Relation of Status to Experience, Perception, and Value," *American Journal of Sociology*, 66, 1–31.

Johansson, J.K., I.A. Ronkainen, and M.R. Czinkota (1994), "Negative Country-of-Origin Effects: The Case of the New Russia," *Journal of International Business Studies*, 25 (10), 157–76.

Keller, K.L. (2002), *Strategic Brand Management: Building, Measuring and Managing Brand Equity*, 2d ed. Upper Saddle River, NJ: Prentice Hall.

Klein, Benjamin and Keith B. Leffler (1981), "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89 (3), 615–39.

Kotler, Philip (1997), *Marketing Management*, 7th ed. Englewood Cliffs, NJ: Prentice Hall.

Lee, C. and R.T. Green (1991), "Cross-Cultural Examination of the Fishbein Behavioral Intentions Model," *Journal of International Business Studies*, 22 (2), 293–305.

Lynn, Michael, George M. Zinkhan, and Judy Harris (1993), "Consumer Tipping: A Cross-Country Study," *Journal of Consumer Research*, 20 (3), 478–89.

Maxwell, Sarah (2001), "An Expanded Price/Brand Effect Model: A Demonstration of Heterogeneity in Global Consumption," *International Marketing Review*, 18 (3), 325–44.

McFadden, D. and K. Train (2000), "Mixed MNL Models for Discrete Response," *Journal of Applied Econometrics*, 15 (5), 447–70.

McGowan, Karen M. and Brenda J. Sternquist (1998), "Dimensions of Price as a Marketing Universal: A Comparison of Japanese and U.S. Consumers," *Journal of International Marketing*, 6 (4), 49–66.

Money, Bruce R., Mary C. Gilly, and John L. Graham (1998), "Explorations of National Culture and Word-of-Mouth Referral Behavior in the Purchase of Industrial Services in the United States and Japan," *Journal of Marketing*, 62 (October), 76–78.

Montgomery, Cynthia A. and Birger Wernerfelt (1992), "Risk Reduction and Umbrella Branding," *Journal of Business*, 65 (1), 31–51.

Nicholls, J.A.F., S. Roslow, and S. Dublish (1997), "Time and Companionship: Key Factors in Hispanic Shopping Behavior," *Journal of Consumer Marketing*, 14 (2–3), 194–205.

Rao, Akshay and Mark Bergen (1992), "Price Premium Variations as a Consequence of Buyers' Lack of Information," *Journal of Consumer Research*, 19 (3), 412–24.

——— and Robert W. Ruekkert (1994), "Brand Alliances as Signals of Product Quality," *Sloan Management Review*, 36 (1), 87–97.

Robinson, Chris (1996), "Asian Culture: The Marketing Consequences," *Journal of the Market Research Society*, 38 (1), 55–63.

Roth, A.E., V. Prasnikar, M. Okuno-Fujiwara, and S. Zamir (1991), "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo: An Experimental Study," *The American Economic Review*, 81 (5), 1068–1095.

Roth, Martin S. (1995), "The Effects of Culture and Socioeconomics on the Performance of Global Image Strategies," *Journal of Marketing Research*, 32 (May), 163–76.

Sekaran, U. (1983), "Methodological and Theoretical Issues and Advancements in Cross-Cultural Research," *Journal of International Business Studies*, 14 (2), 61–73.

Shimp, Terence A. and William O. Bearden (1982), "The Use of Extrinsic Cues to Facilitate Product Adoption," *Journal of Marketing Research*, 19 (May), 229–40.

Spence, Michael (1974), *Market Signaling: Informational Transfer in Hiring and Related Screening Processes*. Cambridge, MA: Harvard University Press.

Steenkamp, Jan-Benedict E.M., Rajeev Batra, and Dana L Alden (2003), "How Perceived Brand Globalness Creates Brand Value," *Journal of International Business Studies*, 34 (1), 53–65.

——— and Hans Baumgartner (1998), "Assessing Measurement Invariance in Cross-National, Consumer Research," *Journal of Consumer Research*, 25 (3), 78–90.

———, Frenkel ter Hofstede, and Michel Wedel (1999), "A Cross-National Investigation into the Individual and National Cultural Antecedents of Consumer Innovativeness," *Journal of Marketing*, 63 (April), 55–69.

Swait, Joffre and Jordan Louviere (1993), "The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models," *Journal of Marketing Research*, 30 (August), 305–314.

Triandis, Harry C. (1995), *Individualism and Collectivism*. Boulder, CO: Westview Press.

——— and Michele J. Gelfand (1998), "Converging Measurement of Horizontal and Vertical Individualism and Collectivism," *Journal of Personality and Social Psychology*, 74 (1), 118–29.

Trompenaars, A. and C. Hampden-Turner (1997), *Riding the Waves of Culture: Understanding Cultural Diversity in Business*. London. N. Brealey Publishers.

Verhage, B.J., U. Yavas, and R.T. Green (1991), "Perceived Risk: A Cross-Cultural Phenomenon," *International Journal of Research in Marketing*, 7 (4), 297–303.

Voich, Dan (1995), *Comparative Empirical Analysis of Cultural Values and Perceptions of Political Economy Issues*. Westport, CT: Praeger.

Wernerfelt, Birger (1988), "Umbrella Branding as a Signal of New Product Quality: An Example of Signaling by Posting a Bond," *Rand Journal of Economics*, 19 (Autumn), 458–66.

Zeithaml, Valarie (1988), "A Consumer's Perceptions of Price, Quality, and Value: A Means–End Model and Synthesis of Evidence," *Journal of Marketing*, 52 (July), 2–22.

Zhou, Kevin Z., Chenting Su, and Yeqing Bao (2002), "A Paradox of Price-Quality and Market Efficiency: A Comparative Study of the U.S. and China Markets," *International Journal of Research in Marketing*, 19 (4), 349–64.

Craig J. Thompson, Aric Rindfleisch, & Zeynep Arsel

Emotional Branding and the Strategic Value of the Doppelgänger Brand Image

Emotional branding is widely heralded as a key to marketing success. However, little attention has been given to the risks posed by this strategy. This article argues that emotional-branding strategies are conducive to the emergence of a doppelgänger brand image, which is defined as a family of disparaging images and meanings about a brand that circulate throughout popular culture. This article's thesis is that a doppelgänger brand image can undermine the perceived authenticity of an emotional-branding story and, thus, the identity value that the brand provides to consumers. The authors discuss how the tenets of emotional branding paradoxically encourage the formation and propagation of doppelgänger brand imagery. This article develops the counterintuitive proposition that rather than merely being a threat to be managed, a doppelgänger brand image can actually benefit a brand by providing early warning signs that an emotional-branding story is beginning to lose its cultural resonance. This article demonstrates these ideas through an analysis of the doppelgänger brand image that is beginning to haunt a paragon of emotional branding—Starbucks. The authors conclude with a discussion of how marketing managers can proactively use the insights gained by analyzing a doppelgänger brand image.

Over the past decade, emotional branding has emerged as a highly influential brand management paradigm (Gobe 2001; Zaltman 2003). Emotional branding is a consumer-centric, relational, and story-driven approach to forging deep and enduring affective bonds between consumers and brands (Roberts 2004). Proponents of emotional branding proclaim that this high degree of consumer passion is seldom, if ever, cultivated through rational arguments about tangible benefits or even appeals to symbolic benefits, such as heightened self-esteem or status (Gobe 2001). Rather, these potent consumer–brand linkages typically emerge when branding strategies use narratives and tactics that demonstrate an empathetic understanding of customers' inspirations, aspirations, and life circumstances and that generate warm feelings of community among brand users (Atkin 2004; Cova and Cova 2002; Fournier 1998; Muniz and Schau 2005).

Among marketing practitioners, this relational, communal, participatory, sensory, and emotive view of consumer–brand relationships is increasingly heralded as a central pillar of market differentiation and sustainable competitive advantage (Atkin 2004; Gobe 2001; Lindstrom 2005; Roberts 2004). Similarly, consumer researchers investigating brand personalities, brand communities, and the meanings that anchor strong consumer–brand relationships arrive at conclusions and recommendations that are consistent with emotional-branding tenets (Aaker 1997; Aaker, Fournier, and Brasel 2004; Brown, Kozinets, and Sherry 2003; Fournier 1998; Kozinets 2001; McAlexander, Schouten, and Koenig 2002; Muniz and O'Guinn 2001). Even brand management paradigms originally premised on cognitive theories of consumer knowledge formation now incorporate many concepts and recommendations that are characteristic of the emotional-branding paradigm (Keller 2003a, b).

However, a seldom-discussed risk of emotional-branding strategies is their potential to expose firms to a particular type of cultural backlash, which we characterize as the "doppelgänger brand image"—that is, a family of disparaging images and stories about a brand that are circulated in popular culture by a loosely organized network of consumers, antibrand activists, bloggers, and opinion leaders in the news and entertainment media. Over time, these brand-focused parodies and criticisms can coalesce into a coherent set of opposing meanings that plague brands that otherwise have attained competitive success through emotional-branding strategies. In effect, a doppelgänger brand image culturally competes with the emotionally resonant meanings that a brand's management attempts to instill through its marketing activities.

From a conventional brand management standpoint, the central question about these brand parodies and social criticisms is whether they actually hurt a brand and whether a firm should take action to combat them (Earle 2002). Indeed, corporations sometimes pursue aggressive legal measures against people who create or distribute parodies of their advertisements or logos. Although this adversarial stance is understandable, we suggest that it is ultimately a

---

Craig J. Thompson is Churchill Professor of Marketing (e-mail: cthompson@bus.wisc.edu), Aric Rindfleisch is Associate Professor of Marketing (and Visiting Professor of Marketing, Tilburg University) (e-mail: aric@bus.wisc.edu), and Zeynep Arsel is a doctoral candidate in Marketing (e-mail zarsel@bus.wisc.edu), School of Business, University of Wisconsin, Madison. The authors thank seminar participants at the University of Arizona and the anonymous *JM* reviewers for their many helpful comments

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

form of brand management myopia that overlooks strategically valuable cultural insights that can be gleaned from an examination of these defamatory representations of the brand.

In this article, we argue that a doppelgänger brand image functions as a diagnostic tool that can reveal latent brand image problems that could eventually blossom into a full-fledged brand image crisis. More specifically, the analysis of a doppelgänger brand image can (1) call attention to cultural contradictions that could potentially undermine the perceived authenticity of a firm's emotional-branding strategy, (2) provide early warning signs that an emotional-branding story is beginning to lose its value as an authenticating narrative for consumers' identity projects, and (3) offer insights into how an emotional-branding strategy can be reconfigured to fit better with changing cultural times and shifting consumer outlooks. We further argue that these diagnostic insights are unlikely to be generated by conventional brand management frameworks.

Our theory-building orientation follows an emerging tradition of studies that use qualitative data and culturally based insights to develop new theoretical knowledge of the symbolic and experiential dimensions of consumer–brand relationships (e.g., Brown, Kozinets, and Sherry 2003; Fournier 1998; Holt 2002; McAlexander, Schouten, and Koenig 2002; Muniz and O'Guinn 2001). In pursuing this agenda, we begin with a brief review of the tenets of emotional branding. We propose that this widely used marketing strategy harbors a set of unintended (and largely unrecognized) risks that predispose brands to the formation of doppelgänger brand images. Next, we show how such unintended consequences emerge in the context of Starbucks's emotional-branding strategy and analyze how these unflattering and disauthenticating meanings motivate consumer brand avoidance. We then discuss general managerial insights that can be garnered by monitoring and analyzing doppelgänger brand images and the broader strategic implications of our research.

The Tenets and Unintended Consequences of Emotional Branding

Emotional branding emerged in the late 1990s as a self-proclaimed corrective to the shortcomings of the conventional benefit-driven approach to branding (see Gobe 2001). Specifically, emotional-branding gurus challenge the benefit-driven approach's fundamental claim that brand managers must establish a clear, consistent, and distinctive benefit position in the mind of the consumer (see Aaker 1996; Ries and Trout 1981). Rather, proponents of emotional-branding argue that a benefit-driven positioning cannot provide an enduring competitive advantage, because it is readily emulated, particularly when the benefits are tied to technological and product design features. They further contend that straightforward benefit appeals are unlikely to break through the clutter of a saturated marketing environment that a plethora of brands are fighting to claim distinctive associations. In this vein, emotional-branding proponents insist that the brand meanings that inspire consumer passion and abiding loyalty are seldom, if ever, attribute-oriented benefits (e.g., Gobe 2001; Mark and Pearson 2001; Roberts 2004).

From an emotional-branding standpoint, brand strategists should focus on telling stories that inspire and captivate consumers. These stories must demonstrate a genuine understanding of consumers' lifestyles, dreams, and goals and compellingly represent how the brand can enrich their lives (Roberts 2004). As Gobe (2002, p. xxi) writes, "consumers today not only want to be romanced by the brands they choose to bring into their lives, they absolutely want to establish a multifaceted holistic relationship with that brand, and this means they expect the brand to play a positive, proactive role in their lives." Thus, the strategic objective of emotional branding is to forge strong and meaningful affective bonds with consumers and, in so doing, become part of their life stories, memories, and an important link in their social networks (Atkin 2004). Over the past few years, many well-known brands have adopted emotional-branding strategies, including Tide, Lexus, Apple, Nike, IBM, Cheerios, McDonald's, and Starbucks (Atkin 2004; Gobe 2001; Roberts 2004).

The benefits of emotional-branding strategies are increasingly being endorsed by the academic marketing community. For example, in a pioneering scholarly article on consumer–brand connections, Fournier (1998) details the ways that consumers can experience brands as relationship partners that help them accomplish personal goals and resolve dilemmas in their everyday lives. Rather than being generated through skillful marketing, Fournier contends that these relational meanings (i.e., love/passion, self-connection, commitment, trust, and intimacy) emerge only when brands become integrated into consumers' lives and identity projects. Several subsequent studies (e.g., Brown, Kozinets, and Sherry 2003; Holt 2002; Kates 2004; Zaltman 2003) have supported the identity significance of emotionally based consumer–brand relationships.

The social dynamics that underlie the emotional-branding paradigm have been most fully explicated in scholarly studies of brand communities, such as the Harley Owners Group and Apple Macintosh enthusiasts (e.g., McAlexander, Schouten, and Koenig 2002; Muniz and O'Guinn 2001; Muniz and Schau 2005; O'Guinn and Muniz 2005). In these communities, brand meanings are collectively forged among a group of consumers. These shared meanings enable consumers from diverse walks of life and different geographic locales to experience deeply satisfying feelings of community and solidarity. It is difficult to find a more striking example of "loyalty beyond reason" (Roberts 2004) than the community of Apple Newton devotees, who have religiously stuck by a brand long abandoned by its manufacturer and who go to arduous extremes to continue using this increasingly antiquated technology (Muniz and Schau 2005).

At present, the conceptual divide between emotional-branding and more traditional benefit-driven branding strategies is rapidly narrowing as emotional-branding principles are being integrated into the benefit-driven brand management paradigm. For example, Keller (2003a, p. 596) commends Zaltman's metaphor elicitation technique, Fournier's brand relationship quality model, and ethnographic research on brand communities for highlighting "intangible aspects of brand knowledge not related to the actual physical product or service specifications per se." The pinnacle of Keller's (2003b, pp. 92–93) customer-based brand equity pyramid is "brand resonance," which refers to "the ultimate relationship and level of attachment that a consumer has with a brand."

Emotional Branding and the Contested Cocreation of Brand Meaning

The tenets of emotional branding are perfectly aligned with the postmodern view that brand meanings are not controlled by managers but rather are cocreated through ongoing interactions among their users (Cova and Cova 2002; Fournier 1998; Muniz and O'Guinn 2001). In the emotional-branding literature, consumers are commonly portrayed as enthusiastic partners who engage in intimate dialogues with other brand users and brand managers to create mutually beneficial, identity-enhancing, community-building, and loyalty-sustaining meanings (Atkin 2004; Gobe 2001; Roberts 2004). The Apple Macintosh and Harley-Davidson brand communities exemplify this model of shared brand ownership. By participating in these collaborative dialogues, brand managers can gain valuable insights into product design improvements and innovations, new product development opportunities, and resonant meanings that can be leveraged in advertising campaigns (Lindstrom 2005; Roberts 2004). In addition, proponents of emotional branding argue that these passionate consumers/owners will often act as brand missionaries, promoting the brand through their own invocative, personalized brand stories (Gladwell 2000; McAlexander, Schouten, and Koenig 2002; Muniz and O'Guinn 2001).

Yet this process of brand-meaning cocreation harbors some infrequently discussed but consequential risks. When a brand is viewed as transgressing or abandoning its emotional-branding promises, formerly loyal customers can create a significant backlash. This pernicious consequence of passionate consumer–brand relationships is vividly illustrated by the travails of Snapple, which was brought to the brink of collapse by a widespread rebellion among its most loyal consumers—energized by the spurned former spokesman Howard Stern, who angrily rechristened the brand "Crapple"—after its new corporate owner, Quaker, attempted to mainstream the brand (Deighton 2002; Holt 2003).

In addition to estranged customers, the media can also circulate brand meanings that are maligning, worrisome, or opprobrious. For example, as Martha Stewart gained celebrity status, she also became the target of an escalating litany of scandalous exposés, unflattering media constructions, and pop culture satires, all of which generated serious brand management headaches for Martha Stewart Living Omnimedia, even before the ImClone scandal (Fournier and Herman 2004). In these alternative cultural constructions, a brand may be used as fodder for humorous parody, cautionary tales about latent lifestyle dangers (e.g., McDonald's vilification in the muckraking documentary *Super Size Me*), or derisive critiques of corporate capitalism (e.g., Nike as the poster child for sweatshop labor). These uncharitable brand meanings and images can quickly spread through multiple channels, such as the news media, entertainment media, e-mail networks, and bloggers (see Klein 1999; Lasn 1999). Consider Procter & Gamble's initial introduction of the fat substitute Olestra, which was effectively sunk by the deadly combination of sensationalized press reports about unpleasant gastrointestinal side effects and the incessant lampooning of this supposed side effect by popular comedians, such as Jay Leno (see Gladwell 2001; Kozinets 2004).

Antibrand activists constitute another vocal group of adversarial brand-meaning creators (see Kozinets and Handelman 2004). These activists, such as Adbusters (www.adbusters.com), are often well-organized entities that try to raise consumer consciousness about what they deem to be the excesses and exploitations of corporate capitalism. When these activists are successful in garnering attention for their antibrand campaigns, a brand's image can take a drastic turn. For example, in the mid-1990s, the cosmetic retailer *The Body Shop* experienced a public relations crisis and lowering of its investment ratings when the popular press brought attention to activists' claims that the company's socially responsible identity was more hype than reality (Atkin 2004; Entine 2003). In effect, these activists were successful in creating a doppelgänger brand image that directly competed with the corporate brand image advanced by *The Body Shop*. A more recent example of potent antibrand activism is offered by Cary and Van Neistat's viral video "iPod's Dirty Secret," which portrays Apple as intentionally exploiting customers by not offering a replacement battery for this widely popular product. Their viral video spread like wildfire across the Internet, and the story was picked up by several mainstream news organizations. The resulting storm of controversy led Apple to offer a replacement battery and, thereby, dramatically alter its marketing strategy (Stuever 2003).

Antibrand activism comes in many forms, including boycotts, demonstrations, organized public relations and informational campaigns, and what has come to be known as "culture jamming" (Kozinets and Handelman 2004; Lasn 1999). Culture jamming refers to a practice of using advertising tools and techniques subversively to cast a critical and often-demonizing light on marketing practices, such as Adbusters' "Joe Chemo" parodies of R.J. Reynolds' controversial Joe Camel campaign and their spoof advertisements featuring jarring juxtapositions between Nike's inspirational advertising rhetoric and stark images of sweatshop laborers (http://www.adbusters.org/spoofads/). Culture jammers tend to target successful iconic brands, such as Nike or Starbucks, because these brands are considered corporate capitalism's most valued (and vulnerable) symbolic assets (Holt 2002; Klein 1999).

Adbusters and like-minded activists characterize themselves as antimarketers who use guerilla tactics to "uncool" iconic brands (Lasn 1999) or, in our terminology, to create doppelgänger brand images that resonate with consumers. As the Internet and e-mail have become ubiquitous aspects of everyday life and as digital-imaging software gains widespread diffusion, the number of brand parodies and biting culture-jamming satires that circulate in cyberspace has mushroomed and are now being recognized as a potentially serious brand image threat (Earle 2002; Walker 2003). Similar to a virulent virus, these alternative, often-unflattering constructions of a brand can move through popular culture with lightening speed.

After a doppelgänger brand image is successfully diffused into popular culture, it presents a competing set of meanings that have the potential to influence consumer beliefs and behavior. We suggest that a doppelgänger brand image can provoke a sense of incongruity between the intimate and inspirational tenor of an emotional-branding strategy and the profit-driven motives of its corporate sponsor. Under such conditions, the authenticity of an emotional-branding story can become suspect, and the identity value that the brand offers to consumers can become greatly reduced (Holt 2002).

The literature on the psychological and cultural dimensions of authenticity is extensive (for more comprehensive reviews, see Campbell 1987; Cross 2000; Grayson and Martinec 2004; MacCannell 1989). A recurrent conclusion of these investigations is that the symbolic divide between the authentic and the inauthentic is a cultural construction rather than an inherent property of objects or actions. Several influential scholars have also noted an enduring cultural tension between the ideal of authenticity and popular conceptions of commercialism, as evinced by conventional laments about artists who compromise their integrity by "selling out" (Lears 1994). Underlying this contrast between the authentic and the commercial is a quasi-religious belief that certain spheres of human existence should remain sacred and distinct from the profane spheres of commerce and profit seeking (Belk, Wallendorf, and Sherry 1989).

The significance that consumers attribute to this ephemeral ideal of authenticity has been documented in several studies (e.g., Arnould and Price 1993; Fournier 1998; Grayson and Martinec 2004; Kozinets 2001; Peñaloza 2000; Price and Arnould 1999; Thompson and Tambyah 1999). Brands that successfully shroud themselves in the cloak of authenticity are able to convey desirable consumer meanings of inner directedness, lack of pretense, and genuine commitment to brand-related activities (see Lewis and Bridger 2000). In pursuit of the authenticity grail, leading brands, such Miller, Burger King, Adidas, and Coke, now routinely use stealth marketing campaigns that are designed to give their brands a more authentic persona through tactics such as the co-option of urban vernaculars and styles and the seamless integration of their brands into the worlds of art and entertainment (Frank 1997; Holt 2002).

However, these tactics are continuously monitored and widely disparaged by antibrand activists as another insidious tactic used by Machiavellian marketers. In a consumer culture marked by pervasive skepticism toward marketing and advertising practices (Walker 2003; Wipperfürth 2005), the authenticity of a brand can be extremely difficult for brand managers to secure and disconcertingly easy for antibrand activists to usurp. When activists propagate a doppelgänger brand image, they are expressing a stinging rebuke to the idea that a corporate-sponsored brand could be a sincere relationship partner or offer genuine emotional sustenance. In the following section, we highlight the emerging doppelgänger brand imagery that is currently challenging the authenticity of Starbucks's emotional-branding claims.

Attacks on Starbucks's Perceived Authenticity

Currently, few brands are riding higher than Starbucks. From its humble origins as a single coffee shop in Seattle a mere 20 years ago, Starbucks has climbed the *Fortune* 500 ladder; it has approximately 7000 stores worldwide and a projected total of 10,000 by the end of 2005 (www.starbucks.com). Starbucks's revenue has climbed annually at an average of 20% per year, and it generated more than $4 billion in gross revenue and nearly $300 million in profit during 2003 (Starbucks Corporation 2003, p. 12), though the chain's same-store sales growth began to slow during 2004 (Ouchi 2004). However, these impressive financial numbers do not fully tell the story of Starbucks's transformative role in popular culture. Starbucks has sparked a dramatic resurgence in coffee consumption in the United States and has changed the leisure practices and daily routines for a broad swath of consumers. The aesthetic look, feel, and sound of Starbucks has influenced the design of retail servicescapes in general (Postrel 2003). Thus, Starbucks has become an important cultural icon, and the "Starbucks revolution" has, in large part, created the burgeoning market for coffee shops and premium-priced gourmet coffee beans (Helliker and Leung 2002).

Given its phenomenal rise, it is no surprise that Starbucks is widely heralded as a shining exemplar of emotional branding and successful brand positioning (e.g., Gobe 2002; Keller 2000; Lewis and Bridger 2000; Roberts 2004; Schmitt 1999; Schmitt and Simonson 1996; Shields 2002). However, not everyone shares this unbridled enthusiasm for the brand. A Web search reveals a staggering number of anti-Starbucks Web sites that present extensively elaborated criticisms of its business practices, scathing diatribes, and manifold culture-jamming ploys. One particular Web posting from an antibrand activist organization in the United Kingdom illustrates the spirit of these anti-Starbucks sentiments (www2.spacehijackers.org/starbucks):

> Starbucks make us feel ill, [w]ith their carefully chosen, subtle (but not too bland) colour scheme. They try to suggest some kind of ethnic, liberal charm. Even the background music has been Starbucked: watered-down, instrumental versions of Hendrix; after all, we wouldn't want to offend. The quirky living room sofas and tables are not so quirky when you realize they are identical to the other thousands of stores, [a]nd those thousands of stores, like a plague infesting our highstreets, are a McDonald's for the new media generation.. . Most of all, we dislike their verminlike spread, subsidizing new stores in order to outprice and close down local rivals, before turning entire towns and cities that horrid green, [b]landing out any local culture and diversity worldwide.

Culture jammers routinely spread anti-Starbucks meanings through cyberspace, using attention-grabbing blends of narrative, satirical images, and animation. One animated

Web parody belittles the pretentiousness of the counter talk that unfolds when a Starbucks' customer orders a drink. Behind the barista is a sign proclaiming "We proudly propagandize Star-Schmucks coffee" and listing ridiculously high prices for standard coffee shop fare. When a trendily dressed customer orders a "vente, white chocolate, mocha thingy," a foul-mouthed cartoon chipmunk (imagine Alvin crossed with Chris Rock) pops up and mercilessly ridicules Starbucks's latte-lingo pretensions and everything else about its customers and marketing tactics (www.illwillpress.com/sml.html). A more scathing viral video likens Starbucks to a Nazi-like empire, steadily and evermore aggressively pursuing world domination (www.albinoblacksheep.com/flash/sbux.php).

Through this gamut of critical parodies, anti-Starbucks activists use humor and satire to assail this brand for many perceived misdeeds, including crushing local competition, degrading the environment, exploiting third-world workers, homogenizing local cultures, and perpetuating the worst excesses of status-conscious consumerism, as evinced by the cultural-jamming takes on its iconic mermaid logo (see Figure 1). Although these brand disparagements may be statistically and factually contestable, they put meanings into cultural circulation that challenge the authenticity of Starbucks's emotional-branding promises.

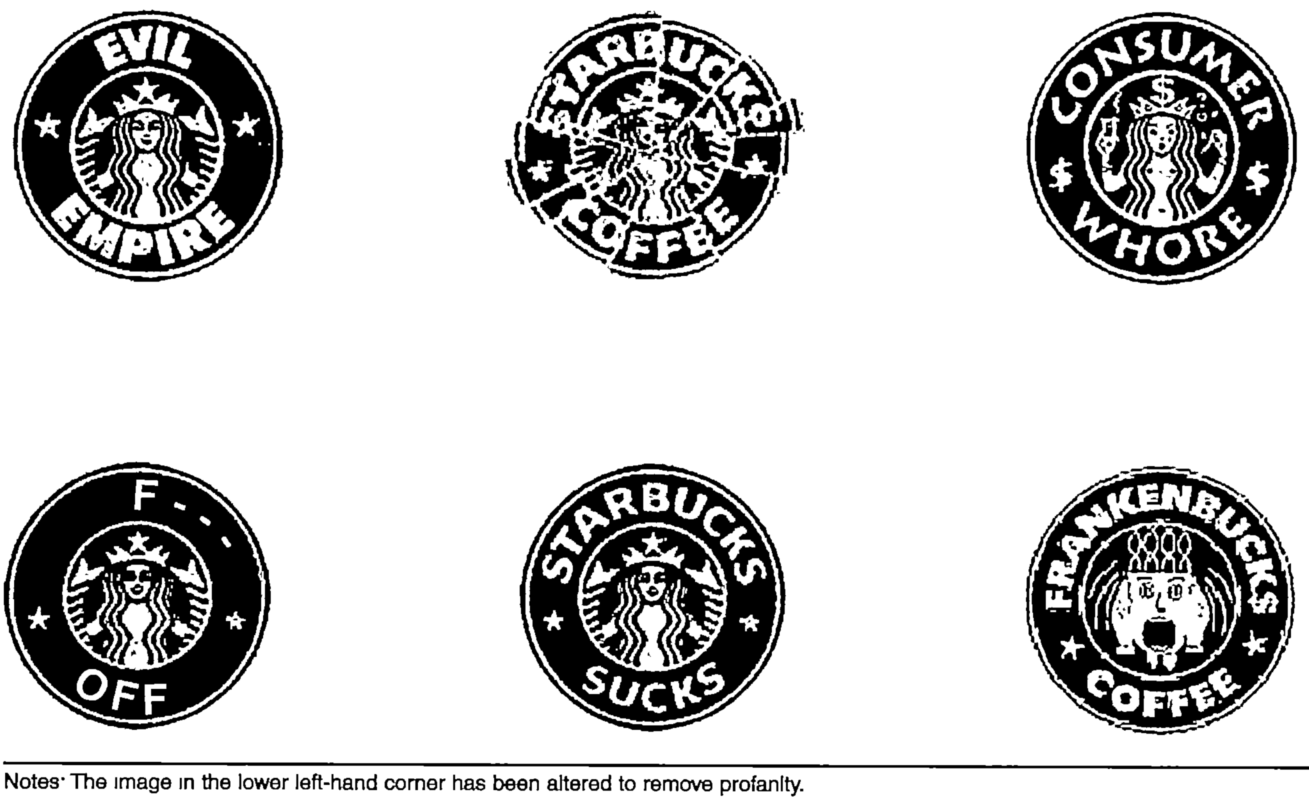
Of course, most consumers are not antibrand activists. However, these culture-jamming criticisms regularly enter into mainstream cultural discourse, such as media satires of Starbucks's breakneck expansion, as exemplified by the *Onion*'s (a satirical newspaper) mock headline "New Starbucks Opens in Rest Room of Existing Starbucks" that has been widely circulated on the Web or the lampooning cartoon presented in Figure 2. Although more sedate in tone, these parodies nonetheless reinforce the doppelgänger image that Starbucks is little more than a profit-driven, marketing affectation and that its customers are either dupes or shameless trend followers.

Much like a conventional brand image, the ultimate meanings of a doppelgänger brand image are forged by the consumers who interpret and use them in the context of their everyday lives. In the next section, we (1) highlight the aspects of Starbucks's doppelgänger brand image that most strongly resonate with our sample of consumers and (2) analyze how these meanings are incorporated into these consumers' preferences, lifestyles, belief systems, and identities. Drawing from these qualitative insights, we then discuss how these seemingly adversarial meanings can be used in a diagnostic fashion to understand, monitor, and proactively manage an emotional-branding strategy more effectively.

An Analysis of Starbucks's Doppelgänger Brand Image

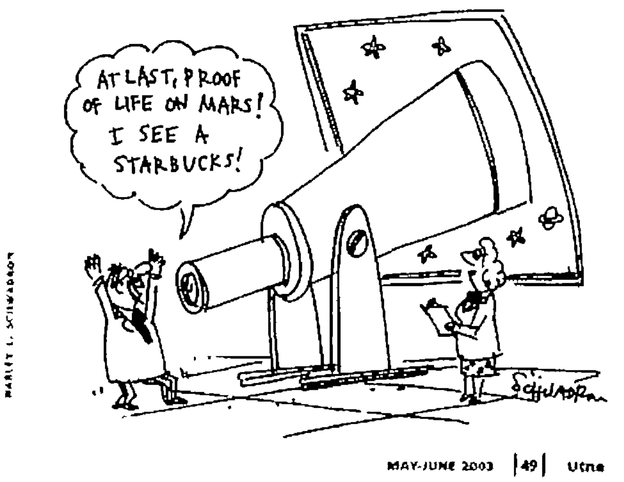
During the course of this study, we collected data over a two-year period, using a multisite approach. Our primary data consist of tape-recorded phenomenological interviews with 36 coffee shop patrons across two locations: a large city and a quintessential "latte town" (see Brooks 2000, pp. 103–109). The interviews totaled more than 60 hours of

FIGURE 1
Culture-Jamming Starbucks's Mermaid Logo



Notes: The image in the lower left-hand corner has been altered to remove profanity.

FIGURE 2
A Popular Culture Parody of Starbucks



Source: From the *Utne Reader*. Reprinted with permission

conversation, which we tape-recorded to produce 600 pages of single-spaced text. These interviews were supplemented by extensive field notes and numerous photographs from on-site observations of patrons across six different local coffee shops.

Interviews were conducted in six different coffee shops, ranging from those that were militantly anti-Starbucks and fashioned around a countercultural motif to those that exhibited a polished aesthetic much like that of Starbucks. This diversity of locales enabled us to accumulate narratives from respondents with a broad range of opinions about the Starbucks brand. Many of our participants also patronized a wide variety of coffee shops, and a few reported that they periodically frequented Starbucks. Thus, our interviews also elicited participants' perceptions and evaluations of the broader coffee shop scene in relation to Starbucks.

Participants were informed that the interviews pertained to their experiences and perceptions of coffee shops and were ensured that their identity would remain confidential. Each interview began with a set of "grand tour" questions (McCracken 1988) about participants' personal backgrounds, interests, and life goals and then focused on their experiences and perceptions of local coffee shops. Interviewers were provided with a small set of predetermined topics to cover (e.g., what participants liked and disliked about their local coffee shop, their perceptions of the coffee shop's regular customers). In keeping with depth interview protocols (e.g., McCracken 1988; Thompson, Locander, and Pollio 1989), however, we allowed participants to guide the flow and content of the discussion to minimize the risk of interviewer-induced biases. These interviews had a conversational quality; interviewers interjected predetermined questions only when breaks in the dialogue arose. Importantly, the interviewers did not introduce the topic of Starbucks. However, in all of our 36 cases, the interviewees raised the distinction between their preferred coffee shop and contrasting doppelgänger meanings attributed to Starbucks. Further probes were then used to elucidate how these consumers understood Starbucks and the kind of identity projected by its brand image.

We interpreted this body of qualitative data using a hermeneutic approach (Thompson 1997). In this approach, provisional understandings are formed, challenged, revised, and further developed through an iterative movement between individual transcripts and the emerging understanding of the entire set of textual data. Thus, each interview was initially treated as a separate idiographic case in which we attempted to uncover the salient meanings and identity projects (e.g., Mick and Buhl 1992) that each participant pursued. As our interpretation unfolded, we attempted to recognize thematic and narrative commonalities that emerged across the data set. Our aim was to identify the most recurrent and robust patterns of underlying cultural meanings that engendered these identified commonalities. This hermeneutic mode of interpretation is premised on the idea that a given consumer is not expressing a strictly subjective viewpoint. Instead, he or she is articulating a culturally shared system of meanings and beliefs, personalized to fit his or her specific life goals and circumstances (Thompson 1997).

Overview of Emergent Themes

> Jack: It's [Starbucks] just a huge corporation. I mean, just a gigantic corporation. They're all over the world. Thousands of outlets, and they come in, run specials, cut prices, and break other coffee shops—just drive them out of business. Because, unfortunately, there's not a lot of loyalty out there as far as consumers are concerned. They're usually going to go where they spend the least amount of money; they're going for the bargains rather than loyalty to local businesspeople.

The doppelgänger meanings plaguing the Starbucks brand are culturally grounded in a particular form of market populism that portrays large corporations as exploitive and rapacious agents whose actions run counter to the best interests of ordinary consumers and small business owners (Cohen 2003; Kelly 2001). However, the accuracy of this populist anticorporatism is widely contested. For example, the two firms most commonly accused of predatory practices, Wal-Mart and Starbucks, counter that they actually improve local economies and create greater opportunities for small businesses. In the case of Starbucks, this anticompetitive charge is particularly ironic because the so-called Starbucks revolution has corresponded to a dramatic rise in the number of locally owned coffee shops (Helliker and Leung 2002).

However, brand image is much more a matter of perceived meaning and cultural mythology (Holt 2003; Mark and Pearson 2001) than an aggregation of verified evidence. Regarding Jack's vignette, for example, there is little evidence that Starbucks, which maintains a premium market position, underprices local competitors. However, Jack's accusation is consistent with the populist idea that large corporations are modern-day robber barons that, as a matter of course, aggressively wield their clout to crush relatively powerless mom-and-pop operations.

The David versus Goliath portrayal of competition between a small, locally owned establishment and a global chain, such as Starbucks, makes for a particularly compelling mythic formulation in which the consumer can also assume heroic qualities. Through patronage of local alternatives, these consumers construe themselves as participating in a dramatic populist struggle against corporate titans, such as Starbucks. Their heroic self-constructions are further buttressed by historically recurrent examples of corporations abusing their economic and political power and putting immediate profits over long-term societal and environmental interests, as exemplified by the well-publicized corporate accounting scandals of 2002.

Among our sample of local coffee shop patrons, the identity value of avoiding the Starbucks brand and the veneration of their preferred local alternatives hinged on two primary motifs, both of which have underlying connections to this brand's lack of perceived authenticity: (1) the cosmopolitan motif and (2) the artisan/owner motif. Through the cosmopolitan motif, local coffee shops are constructed as sites for acquiring authentically distinctive social and aesthetic experiences that are not readily available in corporate-run servicescapes and commercial settings. Through the artisan/owner motif, our participants view local coffee shops as atavistic throwbacks to a bygone era that existed before the marketplace dominance of national chains, when dedicated proprietors put their personal stamps on all facets of the business. Importantly, our data indicate that these authenticating narratives are forged through a critical comparison to Starbucks, which is summarily constructed as little more than a commercialized facsimile of a genuine coffee shop and is perceived as lacking any authentic expression of the cosmopolitan or artisan/owner motifs.

Notably, these two motifs closely mirror Starbucks's highly publicized homage to European cafés (including the attendant ideals of cultural sophistication, worldly engagement, and intellectual conversations) and its related market positioning as a comfortable communal space in which consumers can relax and socialize (Schultz 1997; Shields 2002). Our participants embrace these cosmopolitan and communal motifs as defining features of an authentic coffee shop, and they perceive Starbucks as being unable to deliver on these emotional-branding promises.

A Distinctive Cultural Experience (the Cosmopolitan Motif)

> Scott: I don't like Starbucks, because it seemed like every time I was there, the people have their laptop and then they have their *New York Times* or whatever, their espresso, and it just seems sterile. There was nothing, like here it's a little gothic and a little different. So it seemed too sterile.

Distaste for Starbucks's sterility and boring uniformity is a prominent refrain in our interviews. Our participants widely condemned Starbucks for offering a bland, prepackaged, focus-group-tested consumption experience that caters to bourgeois tastes. In diametric contrast, our participants view local coffee shops as locales for authentic expression of cultural uniqueness and alternative (i.e., noncorporate) sensibilities.

The quest for authentic cultural experiences has been discussed as a prominent motivator of global travels (MacCannell 1989) and excursions to historic sites, museums, and reenactments (Grayson and Martinec 2004; Peñaloza 2000). In an era marked by cultural concerns over marketplace standardization, homogenization, and a growing antipathy toward generic suburban landscapes populated by cookie-cutter homes and strip malls (Oldenburg 1997; Ritzer 1998), the search for authenticity is often coupled with a desire to experience distinctive cultural worlds (Holt 1998; Thompson and Tambyah 1999). This cosmopolitan motif is prominent among our sample of local coffee shop patrons, who consistently characterized an authentic coffee shop as one that provides a spontaneous experience of local color that cannot be attained elsewhere. Importantly, these coveted cosmopolitan experiences are understood through the critical contrast to Starbucks's doppelgänger meaning of being a highly standardized, calculated, and status-conscious servicescape. Consider the following passage:

> Brian: I mean, if you just think of coffee culture, you're thinking about, you know, struggling artists. I hate to stereotype like that, but, I mean, with the local coffee places, you're going to have a little more idiosyncrasy, a little local flavor, color to the place. They might have a specialty kind of menu or something, or they might make [or] do something unique. But, say you're in San Francisco or Seattle, and you see a Starbucks; you more or less know what you're going to get when you walk in the door. This place is kind of an eclectic nature. Yeah, I just like the area. It's kind of a bohemian area of town. It's not like cookie cutter like Starbucks. You don't have to be a certain class to go there; more or less, you can be of any class really. I mean you just, what I like about it is, you never know what kind of character is going to show up here when you're having your coffee.

Through the mundane act of patronizing local coffee shops, our participants can experience themselves as rebels who are consuming against the grain of corporate-dominated mainstream conformity. Through this formulation, their experiences of authentic cultural difference and social diversity afforded by their favorite coffee shops become a resource for cultivating a sense of personal distinctiveness and defiant individuality. Conversely, their diametric contrast to Starbucks's doppelgänger meanings (i.e., corporate conformity and mass-market standardization) functions as a frame of reference that enhances these experiences of authenticity:

> Frank: Well, yeah, a place like this is unique, which I like. It very much caters to, you know, the population around here. It's very, you know, it's comforting to somebody who lives in the suburbs. It's unique, it's got its artsy sort of thing, but as opposed to Starbucks, you know, everywhere it's still the same corporate kind of cookie-cutter type of place that is more, that I don't see the same uniqueness. You know, you can come in here and get a cup of coffee, and it's going to be different than what you get in [name of another local establishment]; it's going to taste a little bit different, but you know, [in] every Starbucks, the food and drinks are pretty much the same. So, I mean, just a place like this has more personality. I wouldn't go to Starbucks. I'll go out of my way to go all the way to the East Side—for me, a drive from here is about six miles. I'd rather do that, and it's just more of a satisfying experience.

Similarly, Ellen echoes Adbusters' critique of mass consumer culture as a stultifying force (Lasn 1999) that inculcates preferences for standardized experiences that are familiar and comforting but ultimately inauthentic:

> Ellen: Well, I actually don't go into coffee shops that are chains, or I avoid them as much as possible. I don't do stores that are chains.
>
> Interviewer: Tell me about that.
>
> Ellen: Well, I don't want to participate in sort of the corporate spread into the world. I have the feeling that people that go to chain places have some sort of sense of their safety in doing that. Like, it's always the same no matter where they go. It actually becomes the culture.
>
> Interviewer : Can you explain that a little more?
>
> Ellen: When you go to Paris or you go someplace out in the world, the tendency is to want to find something that's home, something that you know, and if you walk down the street and find a McDonald's or a Starbucks, you know, this is your home, because these things have become so entrenched in the culture that it's not even conscious. If you go to look for restaurants, to look for places that are locally owned and run, they are not the same all over everywhere and that takes, it feels as if it takes effort. Like people, it takes effort and daring to try something that isn't the same. If you're somebody who's used to going into the same all the time, I think that's a sad comment on our culture.
>
> Interviewer: What are some of the things you think about when choosing a coffee shop?
>
> Ellen: Okay, I check out whether it's a place that has interesting people in it that look like I might want to get to know them. If it offers soy milk (laugh), that's peculiar! If it embraces variety, isn't rushed.... I think that's important too. I like being in a place where people aren't, you know, pounding a cup of coffee and zooming out somewhere. A certain kind of vitality I like as well. I look for where people are alive and active and doing something with their lives.

In this passage, Ellen forges a symbolic contrast between a standardized and mass-marketed consumer culture that caters to a risk-averse and staid clientele and the authentic experiences of local cultural charm that exist outside the sphere of corporate chains. Through her avoidance of chains, Ellen venerates her consumption preferences and lifestyle by pejoratively framing consumers as people who are trapped in the proverbial rat race and who lack time or inclination to live life to its fullest. In sharp symbolic contrast, she aligns herself with an enlightened community that leads authentic and purposeful lives, exhibiting a relaxed sense of joie de vivre and an enthusiasm for new experiences and cosmopolitan pleasures. Much like the value-laden distinction between tourists and travelers (see MacCannell 1989; i.e., "Everyone else is a tourist, but I am a traveler"), a devout preference for local coffee shops and publicly expressed antipathy toward corporate chains, such as Starbucks, can be lionized as a sign of personal virtue and self-directed (rather than marketing-influenced) cosmopolitan tastes.

Beyond the personal significance of these identity-enhancing symbolic distinctions, these consumer narratives also highlight an important cultural aspect about perceptions of authenticity. Prior research has suggested that experiences of servicescape authenticity are steeped in a perceived fidelity to historical circumstances, cultural traditions, or the lives of famous people (Grayson and Martinec 2004; Peñaloza 2000). In the case of the American coffee shop scene, however, these cultural pillars of authenticity carry less symbolic weight. For example, Ellen's preferred coffee shop had been in operation for only three years, so it could not lay claim to some long-standing tradition. In general, local coffee shops tend to be simulations of a romanticized image of the Greenwich Village beatnik/bohemian coffeehouses (circa the 1950s) or left-bank French cafés frequented by beret-wearing, existential-spouting intellectuals. In this context, compelling claims to authenticity must be actively constructed, and the symbolic distinctions forged through Starbucks's doppelgänger brand meanings loom large in this regard. Particularly, the authenticity of local coffee shops' cultural ambiance is contingent on the idea that Starbucks's trademark servicescape is a marketing affectation designed to maximize profitability rather than an authentic, spontaneous expression of local culture:

> Anne: My image of Starbucks is pretty bad. Even though I haven't been there that many times, I think of [it] as more commercial, um, selling out in a way because you think of, like, Starbucks, they find out what people want and the atmosphere and the CDs and the merchandise, and you know, they're doing a really good job with all that, but something about that whole idea.... I just want to rebel against it. It's almost too analytical, where, you know, this place is just more "spiritual," you know, it's more that feeling; it's not as analytical. It's just more emotional or something

Our participants' sharp distinction between authentic coffee shops and crass commercial interests reveals a glaring historical misconception. As McFall (2004) discusses, eighteenth- and nineteenth-century European coffeehouses were protomarketing sites. Not only were they segmented according to different consumer interests, emphasizing poetry/literature, news, storytelling, and gossip, but they also provided an important means to distribute advertisements (with several advertising agencies actually originating from the coffeehouse trade) and commercial gathering points for auctions, sales, and business meetings. However, this historical intertwining of the commercial and the cultural does not easily align with romanticized desires for spontaneous cultural experiences that reveal the defining essence of a community. Through these invidious contrasts to Starbucks, however, local coffee shops can appear to be less driven by marketing considerations and more closely linked to their local community. In other words, our consumers' perceptions of authenticity emerge in relation to their critical constructions of Starbucks and the corporate staging of its servicescape.

An Intimate Commercial Experience (the Artisan/Owner Motif)

> Patrick: *Stammcafé* is a German word for your neighborhood café, where you're a regular, and it's great because it totally says that idea, which is you have your local, local place, you know, your *stammcafé*, the place that you always go to, and then there are "cafés," where you just kind of go and, you know, when you're out with friends or

> if you're in another part of town for a movie or things like that; so sure, why not go there, you know? But the *stammcafé*, you have a special sort of relationship.

Patrick's musings about his *stammcafé* bear a close relation to Oldenburg's (1997) concept of "third places" (i.e., public spaces that function as social gathering points and provide a proverbial home away from home, where a person can experience camaraderie and community). Oldenburg suggests that third places are characterized by a regular clientele who gather to experience a particular form of intimacy in these commercial spaces. This form of intimacy typically does not arise in the formality of the workplace, and it is more relaxing and convivial than the private intimacy of family relationships because it carries fewer interpersonal demands, responsibilities, and obligations. Importantly, the heart and soul of this commercial intimacy is the establishment's proprietor. In the ideal case, patrons believe that they are being personally enriched by their relationship to a unique character (see Oldenburg 1997; Price and Arnould 1999).

This sense of having an authentic connection to an individual proprietor is quite salient to our local coffee shop patrons. However, our interviews and observations suggest that, in general, the servicescapes of American coffee shops do not exhibit the kind of public intimacy and collectively shared conversations that Oldenburg argues are a defining cultural characteristic of third places. The social milieu of the contemporary coffee shop simply does not enjoin such spirited expressions of camaraderie and social bonding. Rather, it tends to foster more individualized experiences of a common connection to the owner/proprietor, who is viewed as the embodiment of the establishment's values and its quirky, endearing qualities.

In this sense, our participants' narratives parallel some findings from Price and Arnould's (1999) analysis of commercial friendships, particularly in regard to the role of an owner's self-revelations in generating a sense of shared intimacy. However, the range of personal information our participants desired is fairly narrow. These consumers want to know just enough about the proprietor to draw inferences about how the coffee shop reflects his or her views and values. Thus, this bounded intimacy enables customers to craft a compelling story about why their preferred coffee shop is truly different and, thus, fundamentally and irrevocably distinct from the corporate-run and standardized Starbucks:

> Interviewer: What is a local coffee shop for you?
>
> Sandra: For me, it would be a coffee shop that isn't national, something that is not a chain. Like this coffee shop, the owner is here most of the time I am here, and I just find that I really respect that, and I have had conversations with the owner before. I know that he used to be, like, in the jewelry business, he left that and opened a coffee shop. So I just find that pretty interesting that the owner frequents the place and works there and has a connection to the customer.
>
> Interviewer: So what is it about knowing the owner or having the owner living in the area that's important to you?
>
> Sandra: I like to support local business. I think that Starbucks, in a way, is very sterile. I mean it doesn't have much character and personality, because it is this mass-produced business, and I think that when you have a situation where there are so many, it becomes more impersonal, and the people who work there don't care as much I would guess. I would think that if it's locally owned, the owner is more involved. He or she would have more connection to the employees, and the employees would probably feel better about working for this person that you see regularly. I mean if you look at this place, it's unique. There is no other place like it in town, there is no other place in the world just like it. But if you look at Starbucks, it's pretty common. It's just not as interesting, I think.

Much like our other participants, Sandra praises local coffee shops for having a one-of-kind quality that emanates directly from the owner's personal touches and that are conspicuously absent from Starbucks's mass-marketed servicescape. Our participants tend to read considerable meaning into their experiences of local charm. They view these distinctive features of the local coffee shops as a kind of self-revelation on the part of the proprietor, who is expressing his or her own personal tastes and values. Furthermore, these various signs of uniqueness are taken as compelling signals that these proprietors are following their own muses rather than calculating what offerings would be most popular to a particular customer segment. Paradoxically, this apparent "take-it-or-leave-it" indifference to customer whims provides these establishments with a particular appeal, particularly in contrast to Starbucks, which is considered a conservative and boring servicescape designed not to offend anyone's sensibilities:

> Kevin: I visited my sister in Oakland; there are three coffee shops on Piedmont Avenue. I've been to all three. One is a place called Peet's coffee; it's like a local that's gone national. Then there's the Starbucks of course, and then there's another little place that I went to that I started to frequent and just loved. It was a little left leaning, kind of different. Their coffee was high quality and stuff, but this little ma-and-pa place served ice cream; they had ice cream as part of their coffee shop—a little different. They had a Ms. Pac-Man video game there, and they also put on the walls photographs of homeless people. They pay homeless people to have Polaroid pictures taken—they would ask for permission of course—and then they gave them, like, $20 or something, and it becomes art on the wall. They weren't afraid to be nontraditional; that's what I loved. Someplace like Starbucks wouldn't dare do that.

In these consumer narratives, local coffee shop proprietors are venerated for being sincerely and deeply committed to their clientele, employees, and the neighborhood at large. In contrast, Starbucks is described as a corporate interloper that wants to reap profits through a mass-produced servicescape and that conspicuously lacks an authentic connection to the local community. Sandra's passage also reiterates the previously discussed heroic framing of the local coffee shop. The proprietor is celebrated for fighting to maintain an invaluable and distinctive community asset against the powerful forces of an impersonal corporate chain. Through this heroic interpretation, Sandra's own patronage of local coffee shops takes on a heightened moral and social significance, representing a means through which she can help ameliorate the blight of standardized corporate servicescapes.

Beyond this immediate sense of personal connection and reciprocal commitment, our participants also portray the proprietors of their favorite coffee shops as devoted artisans, who enact laudable virtues through their business activity, such as educating the palettes of their clientele (in the manner of a wine connoisseur):

> Carie: I love Tina [the owner]. I think, originally, people came into [name of local establishment] for a hit of coffee and a hit of Tina. You know, she's got such great energy. I love to see Tina, and I love the coffee. I love the freshness of their coffee. You know, and I buy their beans. It was really fun when she was first getting into roasting coffee. She had a popcorn popper, like an air popper. Which is very similar to the big roaster that they use. The same technology, so she was experimenting with different beans. She'd bring over, like, little tiny Ziplock bags full of different coffees. So she'd have us try it all the time. At first, it was like a little bit, it was so intense, you're not used to drinking that fresh of a coffee, and then you sort of develop like this, it's like drinking wine. You know, you don't really appreciate, like, a really good wine at first, and then you start to appreciate it. .. She is real connected to the Kaladi Brothers up in Alaska, up in Anchorage, and so it's really fun, because we were up in Anchorage, and you know, Kaladi Brothers was all over the place. Like you'd see a sign, "We serve Kaladi Brothers Coffee." It's the same awesome quality, really consistent, you know, roasting, and everything. My guess is that's her inspiration. And Kaladi Brothers, their main shop is off this road that's difficult to get to; it has no parking, and you know, so the same kind of thing as [her favorite coffee shop's location]. It's not like in the perfect shopping mall and real accessible, so.

This passage exemplifies the shared experience between a proprietor and customers that imbues local coffee shops with an undeniable sense of authenticity. In this case, Carie views herself as intimately involved in the proprietor's (Tina) efforts to perfect a time-honored craft and to cultivate greater consumer appreciation for skillfully roasted coffee. In keeping with the artisan ideal, Tina (the proprietor) has earned a further claim to authenticity through her apprenticeship at Kaladi Brothers Coffee and her choice of an out-of-the-way location for her shop, which contributes to the sense that she is not catering to the masses (a la Starbucks) but only to those who can really appreciate a great cup of coffee (and are willing to make some effort to attain it).

As Brown (2003) suggests, under certain market conditions, inconvenience—the bane of conventional marketing management thought—can actually increase customer loyalty and commitment. Through this nexus of meanings, Carie's patronage of her favorite local coffee shop facilitates an identity project that is particularly salient to consumers with higher levels of cultural capital (Holt 1998). Such consumers strive to demonstrate the distinctiveness and sophistication of their aesthetic tastes by consciously seeking out unique consumer experiences that exist off the proverbial beaten path (Thompson and Tambyah 1999). Starbucks's mass-marketed coffee and its ubiquitous convenience-oriented servicescapes provide a readily recognizable symbol of the consumer mainstream, and thus the brand serves as a symbolic foil for consumers who pursue this sense of cosmopolitan distinctiveness.

Conventional brand management theory holds that, in general, consumers avoid brands whenever negative associations, such as those dogging Starbucks, become strongly established in their minds. Through brand avoidance, consumers circumvent the risk that these stigmatized meanings will sully their public image or self-concept (see Keller 2003b). However, our research provides evidence that consumers do not just avoid doppelgänger brand meanings but rather consciously leverage these disparaging meanings in ways that reinforce an identity-affirming belief that they are more daring and discriminating than the average, mainstream consumer.

Summary

Our cultural analysis suggests that the widespread critiques (by both mass media and antibrand activists) of Starbucks as a rapacious corporate titan are reworked by consumers to create an identity-enhancing morality tale, premised on the consumption ideal of authenticity. The subtext of this authenticating narrative is the David versus Goliath myth, which imbues our participants' experiences of oppositional brand loyalty (see Muniz and O'Guinn 2001) with a particularly strong emotional charge. Our participants perceive local coffee shops as loci of authentic consumption experiences and social relationships. These two themes map onto contradictions that exist between Starbucks's emotional-branding promises and its highly public presence as a global corporate chain. Starbucks's emotional-branding strategy encourages consumers to desire a sense of intimacy and cultural distinctiveness, but paradoxically, it is not as well suited to deliver on these experiential promises as are the myriad local coffee shops that have followed in its frothy corporate wake.

In the following section, we offer some culturally grounded extrapolations beyond our interview data to specify which aspects of Starbucks's emotional-branding strategy contribute to its doppelgänger brand image and render it susceptible to the stigma of inauthenticity. Drawing from this cultural reading, we then illustrate how an analysis of a brand's doppelgänger meanings can be used (1) to diagnose emerging cultural changes that could undermine the authenticity of its emotional-branding story and (2) to gain insights into how this brand story could be altered to avoid this undesirable outcome. This kind of interpretive move is consistent with theoretical arguments that the analyses of consumer narratives gain explanatory power and enhanced managerial relevance when they are supplemented by broader considerations of their supporting social, cultural, and historical contexts (see Arnould and Wallendorf 1994; Holt 2003; Thompson 1997; Zaltman 2003).

Diagnosing Starbucks's Authenticity Dilemma

Starbucks's success is widely attributed to its ability to deliver consistently the communal ambiance, conviviality, and emotional warmth of a quaint European café (Schultz 1997; Shields 2002). According to this explanation, Starbucks has attained market dominance because its carefully designed servicescape is an inviting and comfortable place for consumers to meet, socialize, and enjoy a memorable coffee-drinking experience (Aaker 2004; Keller 2000; Roberts 2004; Schmitt 1999; Schultz 1997). The putative importance of the third-place experiences sits uneasily with tracking studies that indicate that more than 70% of Starbucks's customers are "grab-and-go" types, who are most concerned with speed of service (Moon and Quelch 2004; Stone 2004) and, thus, spend little time imbibing its quaintness and warmth.

This marketplace disparity raises the possibility that this third-place appeal is a peripheral aspect of Starbucks's identity value and emotional appeal. We also note that the disparaging portrayals of the Starbucks brand, whether expressed by our participants or by the gamut of anti-Starbucks Web sites, never question the basic premise that Starbucks is a good place to meet friends and have casual conversations. Rather, challenges to Starbucks's authenticity appear to target a deeper emotional level that cuts to the heart of the brand's identity value.

Taking a cue from Holt's (2004) thesis that brands attain iconic standing when they provide myths that consumers can use to assuage salient threats to their identities, we begin by noting that Starbucks became a cultural icon during this heyday of the mid-1990s new economy and the dot-com boom. These socioeconomic conditions gave rise to a professional class of symbolic workers who were well rewarded for their creative skills. These professionals inhabited organizations that were increasingly decentralized and that reveled in having an informal and playful corporate culture that sharply deviated from bureaucratic norms (Florida 2002).

The social commentator David Brooks (2000) cleverly christened this new professional class as the Bobos, that is, a new breed of professionals who attempted to meld the bourgeois values of hard work, career success, and material affluence with the bohemian values of creativity, expressiveness, anticonformism, and antimaterialism (for a related discussion of how Volkswagen became the automotive brand of choice among the Bobo class, see Holt 2004). Through this postmodern identity, people could be career driven and yet define themselves as anticonformists who live by an edict of creativity and self-expressiveness, and they could lead affluent, brand-conscious lifestyles while viewing themselves as committed antimaterialists.

To maintain a coherent self-identity, Bobos needed a means to assuage the many cultural conflicts posed by these juxtapositions. Thus, for members (and aspirants) of this new class, Starbucks delivered exactly the right kind of mythic story. Through its embodiment of a hip, creative, caring form of capitalism, this brand image erased any sense of cultural incompatibilities between bourgeois and bohemian value systems. Starbucks catered to these new class sensibilities in every aspect of its brand delivery, most particularly its patented decor that emulates not a European café but rather the decor of creative-class office spaces (Florida 2002; Schmitt and Simonson 1996). The symbolic message is that Starbucks is an appropriately informal, creativity conducive space in which the creative class can meld their professional and bohemian identities and directly experience material and symbolic identity validation. Through their patronage of Starbucks, corporate bohemians can reinforce the authenticity of their delicate ideological balancing act and thus allay concerns that they are corporate clones who, in some way, may be culpable for the social and environmental problems widely attributed to corporate capitalism.

The doppelgänger brand meanings we have identified in this study provide some evidence that Starbucks's rapid (and much parodied) market expansion is beginning to undermine this brand's ability to assuage the identity conflicts facing the corporate bohemian class. In effect, Starbucks's doppelgänger brand image unravels its corporate bohemian synthesis and suggests that its bohemian image is a marketing gambit designed to mask its real identity as a corporate titan. This is a potentially severe threat to Starbucks's continued marketing success, because if it loses its aura of authenticity, it can no longer symbolically affirm the Bobo-ish identities of its core customers. In addition, more casual customers may also be more predisposed to question the brand's commitment to social responsibility and the artisan craft of coffee roasting. Such a development would likely make Starbucks vulnerable to competitors in the form of small independents, regional chains, and national competitors, such as Peet's.

The franchise of an iconic brand, such as Starbucks, is typically built around a relatively small segment of extremely loyal customers who validate the authenticity of its emotional-branding story and act as brand evangelists. Through the actions of this devoted core, the brand acquires a heightened cultural cachet that, in turn, attracts a broader segment of peripheral consumers who draw from the brand's symbolic value (Holt 2004; McAlexander, Schouten, and Koenig 2002). As exemplified by the travails of brands such as Snapple and Volkswagen (circa the 1970s and 1980s), when these core customers begin to abandon an iconic brand, it precipitates a downward cultural spiral, which can erode brand equity and market share (see Holt 2004). Thus, a relatively small number of brand avoiders could be the impetus to a cultural tipping point (Gladwell 2000) that weakens Starbucks's brand franchise.

In summary, our analysis indicates that Starbucks's aggressive growth strategy, highly publicized market domination, and omnipresence no longer mesh with its appeal to a countercultural, bohemian sensibility and that this strategy may be compromising the perceived authenticity of its emotional-branding story. However, an analysis of doppelgänger brand meanings can only take brand strategists so far in deciding whether a brand's image should be reconfigured and in identifying what new cultural contradictions should be targeted in an attempt to revise its brand positioning. To accomplish this strategic task, brand strategists need to analyze thoroughly the socioeconomic and cultural conditions currently facing their core customers. For example, the conditions that characterized the "new economy" during Starbucks's rise to iconic status have been fundamentally altered by the puncturing of the dot-com bubble. The current economic climate is colored by the events of September 11 and a heightened sense of economic vulnerability that has arisen from dramatic declines in new job production and other related factors, such as corporations' outsourcing of the new class's "creative work" to lower-wage countries (Pink 2005). By analyzing how these insecurities are represented in popular culture (through books, magazines, films, music, blogs, and other media) and by studying how consumers understand their own lives in relation to them, insights can be gained into resonant life concerns that the Starbucks brand can credibly speak to through its emotional-branding efforts.

As a hypothetical illustration, imagine that follow-up research reveals a growing cultural tension between Americans' historically optimistic view of the future (and the corollary belief that everyone can succeed if they work hard and apply their talents) and a more pessimistic view that economic fate is controlled by forces outside of personal control (and that fate is not always kind or just). This hypothetical cultural contradiction could support a brand mythology that represents Starbucks as a pillar of stability and security in consumers' lives (thus reframing the doppelgänger meanings about Starbucks's "boring standardization" as a positive). Going further, Starbucks could assume the role of an ally or partner that helps consumers take charge of their professional lives and accomplish life goals in the face of random and arbitrary obstacles. Starbucks's professional, corporate aura (again, currently viewed as a negative) and the stimulating properties of its product could lend themselves to this reformulated emotional-branding story.

Implications for Strategic Brand Management

In general, the emotional-branding literature reads like a heroic tale in which intrepid brand managers break with orthodox branding strategies and are rewarded with intensely loyal customers, heightened profits, and a nearly impervious foundation of competitive advantage. These heroic testaments to the power of emotional branding invariably gloss over the potential risks associated with this strategy. Through a detailed examination of the case of Starbucks's emerging doppelgänger brand imagery, we have attempted to address this important issue. In this section, we take a step back from the details of this case and develop a set of initial recommendations that could help brand managers diagnosis pending threats to their emotional-branding investments.

History has shown that dominant brands routinely fall from their stellar heights and struggle to regain their former glory. Disney, McDonald's, AT&T, Levi's, Kodak, and Coca-Cola are just a few examples of market-driven brands that rather suddenly went from paragons of marketing excellence to besieged icons losing ground to competitors (Haig 2003; Holt 2002; Stewart 2002). However, the conventional brand management literature offers little concrete advice on how brand strategists can proactively diagnose the cultural vulnerabilities that could eventually erode their customer-based brand equity. Rather, brand image–related problems are attributed in general to companies failing to follow the tried-and-true principles of sound brand management, such as maintaining consistency and relevance (Keller 2000, 2003a).

Yet the diagnostic value of these normative branding principles is actually quite limited. The consistency principle instructs managers to specify their brand's essence, core personality, and primary benefit carefully and then communicate consistent messages about these essential dimensions. The relevance principle instructs managers to update continuously their brand's user/usage imagery, brand personalities, types of relationships, and promotional themes to keep a brand image fresh, energized, and contemporary (Aaker 2004; Keller 2003a). The somewhat contradictory nature of these two normative principles leaves little doubt that brand managers must undertake a very delicate balancing act. However, these principles beg the important question, How can this balance be effectively attained? To redress this gap, we propose a three-step action plan that can help marketers protect their emotional-branding investments.

Monitor Cultural Cues

First, brand managers should closely monitor popular culture for signs that their brand is beginning to develop a doppelgänger brand image. Here, the critical issue is not how many people embrace these unflattering brand meanings or if they pose an immediate threat to the brand. Indeed, a best-case scenario would be to catch these meanings during their underground phase, analogous to the way that "cool hunters" attempt to discover nascent consumer trends (Gladwell 2000). The relevant concerns at this stage are which aspects of the emotional-branding story are being subverted and what alternative symbolic brand meanings are being created. For example, our analysis of Starbucks suggests that its mass-market ubiquity has undercut an essential component of its emotional-branding story, namely, its ability to act as an authenticating symbol of hip consumerism (see Frank 1997). These cues suggest that Starbucks should consider developing a new emotional-branding story that reframes its corporate connotations in ways that function as a symbolic asset rather than as a source of inauthenticity.

The same medium that facilitates the rapid cultural diffusion of a doppelgänger brand image (i.e., the Internet) can also facilitate the task of cultural monitoring. For example, brand managers should routinely monitor consumer chat rooms, blogs, and Web-based media to gauge the cultural buzz surrounding their marketing campaigns. A more sophisticated approach would be to use systematic netnography techniques (see Kozinets 2002) to assess when an emotional-branding story is beginning to inspire parodies, critics, and other signs of a cultural backlash and to analyze the cultural content of these cultural-jamming efforts. Brand managers should be sensitive to both the frequency and level of intensity of these antibrand sentiments and the specific aspects of their emotional-branding story that are being subverted (and the adversarial meanings that result).

Identify and Track Brand Avoiders

Second, brand managers should identify and study people for whom these doppelgänger meanings resonate and motivate various kinds of brand avoidance. Blogs and antibrand

Web sites may be a useful starting point for locating these people. In addition, customer relationship management systems could be used to identify lapsed customers, who may turn out to be intentionally avoiding a brand because of its doppelgänger imagery. By studying these consumers' viewpoints (we recommend using qualitative methods, such as depth interviews, laddering, ZMET-type collage studies, or even focus groups), managers can assess which doppelgänger brand meanings are resonating with consumers and plumb the identity benefits they gain from this negative imagery.

This second step is a necessary complement to the broader monitoring of cultural cues we outlined in the first step. The ideological concerns that inspire antibrand activists to create and circulate a doppelgänger brand image (and that tend to dominate those representations) may not be that relevant to consumers or that central to their brand avoidance motivations. In the case of Starbucks, for example, antibrand activists widely accuse Starbucks of engaging in a host of predatory business practices. Although the consumers in our study made passing reference to these charges, their primary motivations for avoiding the brand ensued from their perceptions that Starbucks's commercialized servicescape lacked cultural distinctiveness and an authentic personal touch. Thus, brand management decisions based only on an analysis of cultural cues could lead managers to erroneous conclusions about the doppelgänger meanings that pose the greatest threat to their emotional-branding strategies.

Develop and Test a New Emotional-Branding Story

Third, to manage the challenges posed by the emergence of a doppelgänger brand image, we advocate a decidedly proactive approach. Specifically, we believe that brand managers should make adaptations to their emotional-branding story before it goes stale and precipitates a full-blown branding crisis. This proactive orientation carries an unnerving implication: Managers should begin modifying their emotional-branding strategy while it is still accomplishing its marketing objectives. In a dynamic and rapidly changing marketing environment, we believe that it is much better to stay ahead of the cultural curve than to fall behind it and suddenly need to play catch-up with competitors that have more resonant emotional-branding stories.

When confronted by an emotional-branding strategy that is showing blatant signs of diminished marketplace appeal, managers can easily fall into a crisis management panic mode as they grapple with declining sales and lost share. Under such conditions, they may be tempted to delegate much of the strategy-creation process to outside branding consultants in the quest for quick, magic-bullet solutions (Haig 2003). However, this crisis mode can also lead to arbitrary and abrupt shifts in an emotional-branding story that often compound the problem (Keller 2000). For example, if a brand has lost its cultural claim to authenticity, a dramatic shift in the symbolic terms of its emotional-branding strategy could easily be interpreted as further evidence that the brand has no authentic identity other than that dictated by the whims of fashionability.

Accordingly, we suggest that a successful brand image revitalization strategy must fit with prevailing cultural conditions while displaying a clear and comprehensible connection to the emotional-branding story through which the brand made its cultural mark. Because of the amount of corporate transparency now available on the Internet, emotional-branding stories that lack this kind of historical continuity are likely to be perceived as inauthentic (see Holt 2004). A cultural analysis of a doppelgänger brand image can offer valuable insights into how a resonant new emotional-branding story can be culled from the narrative threads of an existing one in a way that maintains this vital sense of historical continuity.

By engaging in the cultural monitoring and brand avoider tracking we discussed previously, brand managers (and their creative consultants) should have a wealth of insights (both cultural and individual) to develop a set of alternative emotional stories that can be tested among consumers. Although most consumers are not particularly good at generating new campaign ideas, particularly ones that tap into emerging cultural trends (see Gladwell 2005; Holt 2003; Zaltman 2003), they are quite astute at assessing the relative degree of emotional connection they experience with different story lines. Again, we recommend the use of qualitative methods, especially rich projective techniques (e.g., Z-MET; Zaltman 2003), as a means to evaluate the emotional resonance (and authenticity value) of alternative story lines. These qualitative insights could also be used to develop quantitative measures of emotional resonance and perceived authenticity (see, e.g., Grayson and Martinec 2004). Furthermore, this approach offers a means to develop a culturally based brand-shielding strategy (see Park, Jaworski, and MacInnis 1986), designed to insulate a brand's symbolic value (e.g., its perceived authenticity) from the devaluation among core customers that can arise when a brand attains broader market appeal.

The implementation of these three steps requires a thorough understanding of a brand's history as well as the cultural meanings and identity projects most salient to its core customers. This type of knowledge is difficult to codify and emerges through a process akin to dynamic puzzle solving. In this puzzle-solving activity, brand strategists must identify a gestalt fit among the target customers' cultural worlds, the properties of the brand, and a credible emotional-branding story that links consumers to the brand. Unlike a conventional puzzle, however, the emotional-branding conundrum keeps changing, and accordingly, emotional-branding strategies must be periodically revised and transformed. In this sense, our proposed guidelines align with Dickson's (1992) competitive rationality thesis that marketing success is dependent on constant monitoring of the dynamic competitive environment and the continual generation of innovative adaptations to these changes.

Conclusions and Future Directions

In this research, we attempted to demonstrate that (1) emotional-branding strategies may unintentionally foster the cultural development of a doppelgänger brand image and (2) careful analysis of this imagery can provide insights into latent branding problems that might not be detected by conventional branding strategy prescriptions. Although we believe that our analysis of Starbucks's emerging doppelgänger imagery provides support for our thesis, we also acknowledge that our research is limited by important boundary conditions and that these conditions represent potential research opportunities.

Our analysis suggests that consumers avoid brands when their emotional-branding promises are viewed as inauthentic and, conversely, that emotional-branding strategies succeed when they can function as an authenticating narrative for consumers' identity projects. These ideas are congruent with research indicating that consumers' most valued brands are those whose symbolic meanings play an important role in their self-conceptions (Fournier 1998; Holt 2002) and with findings that consumers can be unforgiving when trusted brands seem to violate their emotional-branding promises (Aaker, Fournier, and Brasel 2004). However, a lack of perceived authenticity may not be a major stumbling block for brands that do not seek competitive advantage through emotional branding. In addition, brands that lack iconic status (i.e., market challengers or followers) may be less susceptible to attacks on their authenticity because such brands may fly under the radar of antibrand activists. Thus, further research that considers a broader spectrum of brands, consumers, and contexts is necessary to evaluate the robustness of our theory and findings.

REFERENCES

Aaker, David (1996), *Building Strong Brands*. New York: The Free Press.

——— (2004), *Brand Portfolio Management: Creating Relevance, Differentiation, Energy, Leverage, and Clarity*. New York: The Free Press.

Aaker, Jennifer (1997), "Dimensions of Brand Personality," *Journal of Marketing Research*, 34 (August), 347–56.

———, Susan Fournier, and Adam S. Brasel (2004), "When Good Brands Do Bad," *Journal of Consumer Research*, 31 (June), 1–16.

Arnould, Eric J. and Linda L. Price (1993), "River Magic: Extraordinary Experience and the Extended Service Encounter," *Journal of Consumer Research*, 20 (June), 24–45.

——— and Melanie Wallendorf (1994), "Market-Oriented Ethnography: Interpretation Building and Marketing Strategy Formulation," *Journal of Marketing Research*, 31 (November), 484–504.

Atkin, Douglas (2004), *The Culting of Brands: When Customers Become True Believers*. New York: Portfolio.

Belk, Russell W., Melanie Wallendorf, and John F. Sherry (1989), "The Sacred and the Profane in Consumer Behavior: Theodicy on the Odyssey," *Journal of Consumer Research*, 16 (June), 1–39.

Brooks, David (2000), *Bobos in Paradise: The New Upper Class and How They Got There*. New York: Simon & Schuster.

Brown, Stephen (2003), "Torment Your Customers (They'll Love It)," *Harvard Business Review*, 81 (October), 82–88.

———, Robert V. Kozinets, and John F. Sherry Jr. (2003), "Teaching Old Brands New Tricks: Retro Branding and the Revival of Brand Meaning," *Journal of Marketing*, 67 (July), 19–33.

Campbell, Colin (1987), *The Romantic Ethic and the Spirit of Modern Consumerism*. London: Blackwell.

Cohen, Lizabeth (2003), *A Consumers' Republic*. New York: Vintage Books.

Cova, Bernard and Veronique Cova (2002), "The Tribalization of Society and Its Impact on the Conduct of Marketing," *European Journal of Marketing*, 36 (5–6), 595–620.

Cross, Gary (2000), *An All-Consuming Century: Why Commercialism Won Out in America*. New York: Columbia University Press.

Deighton, John (2002), "How Snapple Got Its Juice Back," *Harvard Business Review*, 80 (January), 47–52.

Dickson, Peter (1992), "Toward a General Theory of Competitive Rationality," *Journal of Marketing*, 56 (January), 69–83.

Earle, Carey (2002), "FWD. This Made Me Laugh: How Viral Brand Parodies Impact Your Brand," (accessed September 1, 2005), [available at http://www.brandchannel.com/images/papers/FWD.pdf].

Entine, Jon (2003), "The Myth of Social Investing: A Critique of Its Practices and Consequences for Corporate Social Performance Research," *Organization & Environment*, 16 (September), 352–68.

Florida, Richard (2002), *The Rise of the Creative Class: And How It's Transforming Work, Leisure, Community and Everyday Life*. New York: Basic Books.

Fournier, Susan (1998), "Consumers and Their Brands: Developing Relationship Theory in Consumer Research," *Journal of Consumer Research*, 24 (March), 343–74.

——— and Kerry Herman (2004), "Taking Stock in Martha Stewart," working paper, Tuck School of Business, Dartmouth College.

Frank, Thomas (1997), *The Conquest of Cool: Business Culture, Counterculture, and Hip Consumerism*. Chicago: University of Chicago Press.

Gladwell, Malcolm (2000), *The Tipping Point: How Little Things Can Make a Big Difference*. Boston: Little, Brown and Company.

——— (2001), "The Trouble with Fries: Fast Food Is Killing Us. Can It Be Fixed?" *The New Yorker*, (March 5), 52–57.

——— (2005), *Blink: The Power of Thinking Without Thinking*. New York: Little, Brown and Company.

Gobe, Marc (2001), *Emotional Branding: The New Paradigm for Connecting Brands to People*. New York: Allworth Press.

——— (2002), *Citizen Brand*. New York: Allworth Press.

Grayson, Kent and Radan Martinec (2004), "Consumer Perceptions of Iconicity and Indexicality and Their Influence on Assessments of Authentic Market Offerings," *Journal of Consumer Research*, 31 (September), 296–313

Haig, Matthew (2003), *Brand Failures: The Truth About the 100 Biggest Branding Mistakes of All Time*. Sterling, VA: Kogan.

Helliker, Kevin and Shirley Leung (2002), "Counting Beans: Despite the Jitters, Most Coffeehouses Survive Starbucks," *The Wall Street Journal*, (September 24), A1, A11.

Holt, Douglas B. (1998), "Does Cultural Capital Structure American Consumption?" *Journal of Consumer Research*, 25 (June), 1–26.

——— (2002), "Why Do Brands Cause Trouble? A Dialectical Theory of Consumer Culture and Branding," *Journal of Consumer Research*, 29 (June), 70–90.

——— (2003), "What Becomes an Icon Most?" *Harvard Business Review*, 81 (March), 43–49

——— (2004), *How Brands Become Icons: The Principles of Cultural Branding*. Cambridge, MA· Harvard Business School Press.

Kates, Stephen (2004) "The Dynamics of Brand Legitimacy: An Interpretive Study in the Gay Men's Community," *Journal of Consumer Research*, 31 (September), 455–64.

Keller, Kevin Lane (2000), "The Brand Report Card," *Harvard Business Review*, 78 (January–February), 147–54.

——— (2003a), "Brand Synthesis: The Multidimensionality of Consumer Knowledge," *Journal of Consumer Research*, 29 (March), 595–600.

——— (2003b), *Strategic Brand Management: Building, Measuring, and Managing Brand Equity*, 2d ed. Upper Saddle River, NJ: Prentice Hall.

Kelly, Marjorie (2001), *The Divine Right of Capital: Dethroning the Corporate Aristocracy*. San Francisco: Berrett-Koehler.

Klein, Naomi (1999), *No Logo: Taking Aim at the Brand Bullies*. New York: Picador.

Kozinets, Robert V. (2001), "Utopian Enterprise: Articulating the Meaning of Star Trek's Culture of Consumption," *Journal of Consumer Research*, 28 (June), 67–89.

——— (2002), "The Field Behind the Screen: Using Netnography for Marketing Research in Online Communities," *Journal of Marketing Research*, 39 (February), 61–82.

——— (2004), "Procter & Gamble's Olestra: Dealing with New Product Introduction Issues," University of Wisconsin School of Business case. Madison: University of Wisconsin.

——— and Jay M. Handelman (2004), "Adversaries of Consumption: Consumer Movements, Activism, and Ideology," *Journal of Consumer Research*, 31 (December), 691–704.

Lasn, Kalle (1999), *Culture Jam: The Uncooling of America*. New York: William Morrow.

Lears, Jackson (1994), *Fables of Abundance: A Cultural History of Advertising in America*. New York: Basic Books.

Lewis, David and Darren Bridger (2000), *The Soul of the New Consumer: Authenticity—What We Buy and Why in the New Economy*. London: Nicholas Brealey Publishing.

Lindstrom, Martin (2005), *Brand Sense: Build Powerful Brands Through Touch, Taste, Smell, Sight, and Sound*. New York: The Free Press.

MacCannell, Dean, (1989), *The Tourist: A New Theory of the Leisure Class*, 2d ed. New York: Schocken Books.

Mark, Margaret and Carol Pearson (2001), *The Hero and the Outlaw: Building Extraordinary Brands Through the Power of Archetypes*. New York: McGraw-Hill.

McAlexander, James H., John W. Schouten, and Harold F. Koenig (2002), "Building Brand Community," *Journal of Marketing*, 66 (January) 38–54.

McCracken, Grant (1988), *The Long Interview* Thousand Oaks, CA: Sage Publications.

McFall, Liz (2004), "The Language of the Walls: Putting Promotional Saturation in Historical Context," *Consumption Markets and Culture*, 7 (June), 107–128.

Mick, David Glen and Claus Buhl (1992), "A Meaning-Based Model of Advertising Experiences," *Journal of Consumer Research*, 19 (December), 317–338.

Moon, Youngme and John Quelch (2004), "Starbucks: Delivering Customer Service," Harvard Business School Case No. 9-506-016. Cambridge, MA: Harvard Business School Press.

Muniz, Albert M. and Thomas C. O'Guinn (2001), "Brand Community," *Journal of Consumer Research*, 27 (March), 412–32.

——— and Hope Jensen Schau (2005), "Religiosity in the Abandoned Apple Newton Brand Community," *Journal of Consumer Research*, 31 (March), 737–47.

O'Guinn, Thomas C. and Albert M. Muniz (2005), "Communal Consumption and the Brand," in *Inside Consumption: Frontiers of Research on Consumer Motives, Goals, and Desires*, S. Ratneshwar and David Glen Mick, eds. New York: Routledge, 252–72.

Oldenburg, Ray (1997), *The Great Good Place*, 2d ed. New York: Marlowe & Company.

Ouchi, Monica Soto (2004), "Starbucks Sales Cool Off," *Seattle Times*, (August 27), (accessed September 1, 2005), [available at http://seattletimes.nwsource.com/html/businesstechnology/2002016813_starbucks27.html].

Park, C. Whan, Bernard J. Jaworski, and Deborah J. MacInnis (1986), "Strategic Brand Concept–Image Management," *Journal of Marketing*, 50 (October), 135–45.

Peñaloza, Lisa (2000), "The Commodification of the American West: Marketers' Production of Cultural Meanings at a Trade Show," *Journal of Marketing*, 64 (October), 82–109.

Pink, Daniel H. (2005), *A Whole New Mind*. New York: Riverhead Books.

Postrel, Virginia (2003), *The Substance of Style*. New York: HarperCollins.

Price, Linda L. and Eric J. Arnould (1999), "Commercial Friendships: Service Provider–Client Relationships in Context," *Journal of Marketing*, 63 (October), 38–56.

Ries, Al and Jack Trout (1981), *Positioning: The Battle for Your Mind*. New York: McGraw-Hill.

Ritzer, George (1998), *The McDonaldization Thesis: Explorations and Extensions*. Thousand Oaks, CA: Sage Publications.

Roberts, Kevin (2004), *Lovemarks. The Future Beyond Brands*. New York: Powerhouse Books.

Schmitt, Bernd H. (1999), *Experiential Marketing: How to Get Customers to Sense, Feel, Think, Act, and Relate to Your Company and Brands*. New York: The Free Press

——— and Alex Simonson (1996), *Marketing Aesthetics: The Strategic Management of Brands, Identity, and Image*. New York: The Free Press.

Schultz Howard (1997), *Pour Your Heart into It: How Starbucks Built a Company One Cup at a Time*. New York: Hyperion.

Shields, David (2002), "The Capitalist Communitarian," *The New York Times Magazine*, (March 24), 28–31.

Starbucks Corporation (2003), *Starbucks Annual Fiscal Report*. Seattle: Starbucks Corporation.

Stewart, David W. (2002), "Getting Published: Reflections of an Old Editor," *Journal of Marketing*, 66 (October), 1–6.

Stone, Brad (2004), "Grande Plans: Yes, It Can Seem Like There's a Starbucks on Every Corner, but Howard Schultz Says His Company Is Just Warming Up," *Newsweek Business*, (October 4), (accessed September 1, 2005), [available at http://www.msnbc.msn.com/id/6100243/site/newsweek/].

Stuever, Hank (2003), "Battery and Assault: When His iPod Died, This Music Lover Tackled Apple. Stay Tuned," *The Washington Post*, (December 20), C01.

Thompson, Craig J. (1997) "Interpreting Consumers: A Hermeneutical Framework for Deriving Marketing Insights from the Texts of Consumers' Consumption Stories," *Journal of Marketing Research*, 34 (November), 438–55.

———, William B. Locander, and Howard R. Pollio (1989), "Putting Consumer Experience Back into Consumer Research: The Philosophy and Method of Existential-Phenomenology," *Journal of Consumer Research*, 16 (September), 133–47.

——— and Siok Kuan Tambyah (1999), "Trying to Be Cosmopolitan," *Journal of Consumer Research*, 26 (December), 214–41.

Walker, Rob (2003), "The Marketing of No Marketing," *The New York Times*, (June 22), (accessed September 1, 2005), [available at http://www.times.com/].

Wipperfurth, Alex (2005) *Brand Hijack: Marketing Without Marketing*. New York. Portfolio.

Zaltman, Gerald (2003), *How Customers Think: Essential Insights into the Mind of the Market*. Cambridge, MA: Harvard University Press.

Martin Spann & Gerard J. Tellis

Does the Internet Promote Better Consumer Decisions? The Case of Name-Your-Own-Price Auctions

The Internet has enabled consumers to make more informed decisions more conveniently with apparently more efficient price-clearing mechanisms than was available before its advent. One such mechanism is the name-your-own-price auction. The authors study the extent to which decisions made in such an auction are rational in relation to an economic model. The results from two independent markets show that a large number of consumers do not exhibit rational decision making.

The Internet has made it easier for sellers to inform consumers of their offers and for consumers to respond to those offers than has been possible in traditional retail channels (Bakos 1997; Brynjolfsson and Smith 2001; Geoffrion and Krishnan 2003). Several retail forms have evolved on the Internet, such as single-store offers, comparison sites with offers across multiple stores, and various types of auctions. One type of auction is the name-your-own-price mechanism (e.g., www.priceline.com), in which consumers bid for a product on sale at an unrevealed retailer's price. At Priceline, consumers are allowed to bid only once within seven days for a specific product (e.g., a flight); however, they can partially circumvent the time constraint (Fay 2004). Other name-your-own-price retailers have extended the model to allow multiple bidding for the same product.

Ideally, auctions enable a better market-clearing price between a buyer and a seller, assuming that both are rational. Sellers hope to extract the maximum buyer surplus by completing a transaction at the highest price that a buyer is willing to pay above their own marginal cost for that item. Buyers hope that they can obtain the lowest price a marketer is willing to take below their own reservation price. However, because buyers do not know the sellers' marginal cost, the only way a buyer can obtain such a good deal is to begin with a low price and raise it in increments up to his or her reservation price until the seller agrees to the price. Economic theory suggests the specific sequence of prices that a rational, price-minimizing consumer should adopt. However, to do so, buyers need time, thoughtfulness, and planning, all features that the Internet is supposed to provide. The specific issue we address in this article is whether consumers' actual bids fit the pattern suggested by economic theory in the case of name-your-own-price auctions. In general, we ascertain the extent to which shopping on the Internet facilitates consumers' rational, price-minimizing decisions.

The literature on name-your-own-price auctions is sparse. Some studies focus on the specific design of Priceline but do not provide empirical data (Chernev 2003; Ding et al. 2005; Fay 2004). A few studies analyze consumer characteristics on the basis of data on such auctions, but they do not examine whether consumer behavior is rational (Hann and Terwiesch 2003; Spann, Skiera, and Schäfers 2004). The current study analyzes the empirical behavior of consumers and compares it with the economic norm to ascertain the extent to which it is rational.

We conduct this analysis in three steps. First, we analyze whether observed bidding patterns follow the normative patterns. Second, we analyze what factors explain deviations from rational behavior. Third, we analyze the effect of these possible influencers of rational behavior on the value of consumers' bids. We apply our analysis to two large data sets of consumers' bids using a name-your-own-price auction: (1) a retailer selling flights from one country to various international and domestic destinations and (2) a low-cost airline selling its own flights within Europe.

We organized the remainder of this article as follows: In the next section, we describe the expected behavior of consumers at the name-your-own-price auction. Then, we describe the data and our analysis and present the results. Finally, we discuss the implications of our findings.

Expected Behavior at Name-Your-Own-Price Auctions

Name-Your-Own-Price Auction

In a name-your-own-price auction, any consumer who bids above a seller's unrevealed threshold price receives the product at the price of his or her bid. In case of limited availability, consumers who are the first to bid above the threshold are served first. In contrast, a standard auction

---

Martin Spann is Professor of Marketing and Innovation, School of Business and Economics, University of Passau, Germany (e-mail: spann@spann.de). Gerard J. Tellis is Professor of Marketing, Director of the Center for Global Innovation, and Neely Chair of American Enterprise, Marshall School of Business, University of Southern California, Los Angeles (e-mail: tellis@marshall.usc.edu). The authors gratefully acknowledge many helpful comments from the three anonymous *JM* reviewers and from Arina Soukhoroukova, Bernd Skiera, Martin Bernhardt, Björn Schäfers, and Oliver Hinz.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

determines the winning bidder as the one who places the highest bid (if bidding to buy) or the lowest bid (if bidding to sell) among rival bids.

Two broad types of name-your-own-price auctions are prevalent. In one type, firms more or less perfectly limit consumers to a single bid for a specific product. An example of such an auction is that of Priceline (Fay 2004). In the other type, firms allow consumers to bid again if their previous bid has been rejected. Several examples of the latter are available in Europe. In the latter type, if the seller accepts a buyer's bid, a purchase occurs because consumers' bids are binding (e.g., Priceline requires consumers to provide their credit card information before bidding). If a consumer's bid is unsuccessful, he or she can either withdraw or submit another bid (see Figure 1). Rejected bids provide some information about buyers and sellers. By bidding, consumers provide the lower limit of the range of their reservation price at that moment. In rejecting the bid, sellers indicate the lower limit of the range of their threshold price at that moment. According to economic theory, the threshold price should represent the sellers marginal cost for the product unless the seller hopes to attract consumers with higher prices.

Buyers' Optimal Behavior

How should buyers bid in a name-your-own-price auction such as that which we previously described? Economic theory provides a framework to answer this question, assuming that consumers are rational and want to minimize the prices they pay. Spann, Skiera, and Schäfers (2004) develop a model that arrives at the consumer's optimal strategy in terms of the optimal number of bids and the optimal prices of those bids. They apply the model to impute consumers' willingness to pay and bidding costs. We used their model to derive the optimal pattern of bid prices. Here, we describe the economic intuition of the model and derive the optimal pattern of bid prices. (Readers may go to the original article for details of the model development.)

In this model, consumers maximize their surplus, which is the difference between their reservation price and the price they pay. The latter is the amount of their bid if it is successful (i.e., above the firm's threshold price, which is unknown to consumers). Because this threshold price, and thus the success of a bid, is uncertain to consumers, they maximize their expected surplus (i.e., surplus × probability of success). Therefore, consumers need to form a belief about the probability distribution of the unknown threshold price. By submitting a bid, consumers accrue costs that reduce their surplus. The costs of a bid include the cost of thinking to determine the optimal amount of the bid, the transaction costs for the time and trouble of bidding, and the costs of waiting for the acceptance or rejection of a bid (Shugan 1980; Tellis 1986).

Consumers influence their surplus and the probability that a bid is successful by the size of their bid. On the one hand, the consumer can increase the probability that the bid is successful (i.e., above the seller's unrevealed threshold price) by increasing the amount of the bid. On the other hand, a higher bid leads to a lower consumer surplus in the case of a successful bid. Thus, the consumer must solve this trade-off between increasing consumer surplus and decreasing the probability of a successful bid. Consumers do not bid if they expect a negative surplus.

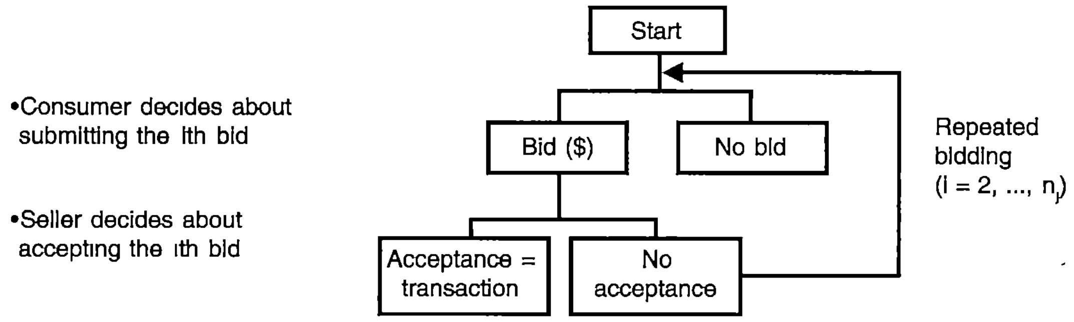
The consumer has the possibility of rebidding if a bid is not successful. In this case, consumers update their belief about the distribution of the threshold price because an unsuccessful bid signals a threshold price that is greater than the bid. Thus, the updated distribution for the threshold price is left-truncated at the value of the unsuccessful previous bid. The optimal bid price of a repeated bid is then determined on the basis of this updated belief (for the formal model, see Appendix A).

The optimal pattern of bid prices has three norms (for the proof, see Appendix B). First, consumers should begin with the lowest price that does not seem completely unreasonable given their reservation price and belief about the threshold price. The reason is that consumers do not know how many seats are left, under what duress the seller operates, and what the seller's marginal cost is. A reasonably low price can gather information about these factors.

Second, if their bids are rejected, buyers should increase them up to their reservation price, until the seller accepts the bid. The reason for increasing the bid is to offer a price that is more acceptable to sellers, as long as it is below the buyer's reservation price.

Third, if not under pressure of time, buyers should increase their bids in decreasing size of increments. The reasoning behind this third norm is the most difficult to

FIGURE 1
Multiple Bidding at a Name-Your-Own-Price Retailer



appreciate. The logic is as follows: Costs accruing with each bid reduce consumers' surplus. Given higher costs, the optimal strategy is to reduce the number of bids necessary to be successful. Consumers increase the probability of each bid being successful by bidding higher than their prior bid. However, at the same time, as they bid closer to their reservation price and the seller's unknown threshold price, they do not want to bid higher than is necessary to avoid giving up too much of their surplus to the seller. Therefore, they should increase their bid in smaller increments.

If we assume rational, price-minimizing buyers, the optimal bidding strategy is characterized with monotonically increasing bids, which have decreasing increments as long as consumers accrue bidding costs. We refer to such bidding sequences as being "strongly rational." If bidding costs are zero, consumers' bids monotonically increase at a constant rate. However, such costs are unlikely to be zero (e.g., Hann and Terwiesch 2003; Spann, Skiera, and Schäfers 2004). If consumers are under pressure or if they panic about not winning a bid, they may bid with increasing increments. We refer to this case as being "weakly rational."

Supply constraints, such as limited seats, can affect the probability of whether a consumer's bid is successful. Rational consumers would make estimates of these supply constraints as best they can before their first bid. They incorporate this risk of a supply constraint in their belief about the probability of winning the bid and the distribution of the unknown threshold price. They estimate the threshold price to be higher than when there are no constraints. Because rational consumers determine *ex ante* their optimal bidding behavior by simultaneously determining the optimal number of bids and optimal bid prices, the belief about supply constraints is already incorporated into their optimal bidding strategy. Next, we examine real markets to determine whether and to what extent observed buyer behavior conforms to the normative behavior we described previously (again, assuming rational, price-minimizing buyers).

Data

Study 1: Name-Your-Own-Price Retailer

Our first study examines the data of a name-your-own-price retailer based in Germany that sells airline tickets for various airlines and allows multiple bidding. Our data set consists of all bids at the name-your-own-price retailer for a period of 11 months between February 2000 and December 2000. We have 6539 bidding sequences and a total of 12,999 individual bids for flights from Germany to 86 different international and domestic destinations. A bidding sequence consists of the bids of a specific consumer for a specific destination. Table 1 shows the distribution of bidding sequences by bid length and the distribution of consumers by destinations for which they bid: 78% of all bidding sequences consist of only one or two bids, and 88% of all consumers bid for only one destination. Table 2 shows

TABLE 1
Description of Data from Name-Your-Own-Price Retailer

Distribution of Bidding Sequences by Bid Length		Distribution of Consumers by Destinations	
Bid Length	**Number of Sequences**	**Destinations Bid**[a]	**Number of Consumers**
1	3647	1	4965
2	1477	2	492
3	634	3	110
4	347	4	25
5	173	5	12
6	89	6	12
>6	172	7	4
All	6539	All	5620

[a] Destinations for which a consumer bids.

TABLE 2
Number of Destinations by Geographic Region (Name-Your-Own-Price Retailer)

Geographic Region	Destinations	Sequences	Proportion (%)[a]
Germany (domestic)	7	1160	18
Europe	32	3397	52
North America	21	1240	19
Rest of world	26	742	11
All	86	6539	100

[a] Proportion of sequences per geographic region according to all sequences.

the number of destinations by geographic region: Fewer than one fifth of all bidding sequences are for domestic flights, but a majority of sequences are for flights within Europe.

Study 2: Name-Your-Own-Price Low-Cost Airline

Our second data set consists of the application of the name-your-own-price auction during a promotional offering by a European low-cost airline. Low-cost airlines (e.g., Southwest, easyJet, JetBlue) differ from traditional airlines in several ways. First, they offer "no-frills" service at low prices. Second, they treat each flight as a separate product; that is, they do not apply any booking restrictions, such as a maximum stay restriction between the outward and the return flight. Third, they usually do not serve transcontinental destinations. Our data set consists of all bids for flights on this airline from a name-your-own-price auction during a promotional offering (from Friday, September 17, to Sunday, September 19, 2004). Consumers could bid for 14 different flight destinations in Europe on 30 different dates (between the end of September 2004 and the end of October 2004). We have 3598 bidding sequences and 12,280 individual bids for these destinations. Table 3 shows the distribution of bidding sequence by bid length and the distribution of consumers by destinations for which they bid: 53% of all bidding sequences consist of only one or two bids, and 80% of all consumers bid for only one destination.

Modeling Bidding Behavior

We analyze whether empirically observed patterns match the normative pattern of behavior (assuming rational, price-minimizing buyers). We do so through a hierarchical model that consists of two stages (e.g., Chandy et al. 2001). In the first stage, we estimate each consumer's pattern of bidding. In the second stage, we ascertain whether the patterns of bidding estimated in the first stage vary by characteristics of consumers and their bids.

First-Stage Analysis

Classification of bidding patterns. We consider only bidding sequences that are longer than two bids for a single itinerary. To capture various bidding patterns parsimoniously, we fit the following quadratic function to each sequence of bids:

$$(1)\quad B_{k,i} = \alpha_{1k} \times I_{k,i}^2 + \alpha_{2k} \times I_{k,i} + \alpha_{3k} + \varepsilon_{k,i} \qquad (I_{k,i} > 0),$$

where $B_{k,i}$ is the bid value; $I_{k,i}$ is the number of a bid in a kth bidding sequence; α_{1k}, α_{2k}, and α_{3k} are parameters that capture the shape of the quadratic function; and $\varepsilon_{k,i}$ are residuals, assumed to follow a normal distribution identically and independently. We estimate Equation 1 by minimizing the sum of squared residuals for each individual sequence to capture consumer heterogeneity.

On the basis of the estimated parameter values α_{1k} and α_{2k}, the first couple of bids, and the last bid number $I_{k,max}$ of each (kth) bidding sequence, we classified individual sequences into six different bidding patterns (see Figure 2). The constant term α_{3k} is irrelevant for classification. We begin by explaining the identification of the two simplest patterns (Pattern 2 and 6) and then move on to the more complex patterns.

Pattern 2 reflects linearly increasing bids. A sequence follows Pattern 2 if the quadratic term has a parameter of zero ($\alpha_{1k} = 0$) and the linear term has a parameter of positive value ($\alpha_{2k} > 0$).

Pattern 6 reflects constant or linearly decreasing bids. A sequence follows Pattern 6 if the quadratic term has a parameter of zero ($\alpha_{1k} = 0$) and the linear term has a parameter of negative value or zero ($\alpha_{2k} \leq 0$).

Pattern 1 reflects increasing bids at a decreasing rate; the last bid still is the highest in the sequence (i.e., only the upward portion of an inverted U shape). A sequence follows Pattern 1 if the parameter of the linear term is positive ($\alpha_{2k} > 0$), the parameter of the quadratic term is negative ($\alpha_{1k} < 0$), and the value of the last bid of the sequence is the highest ($B_{k,Imax} > B_{k,Imax-1}$). Appendix C shows that these conditions are met when

$$(2)\qquad 0 > \alpha_{1k} > \frac{-\alpha_{2k}}{2 \times I_{k,max} - 1}.$$

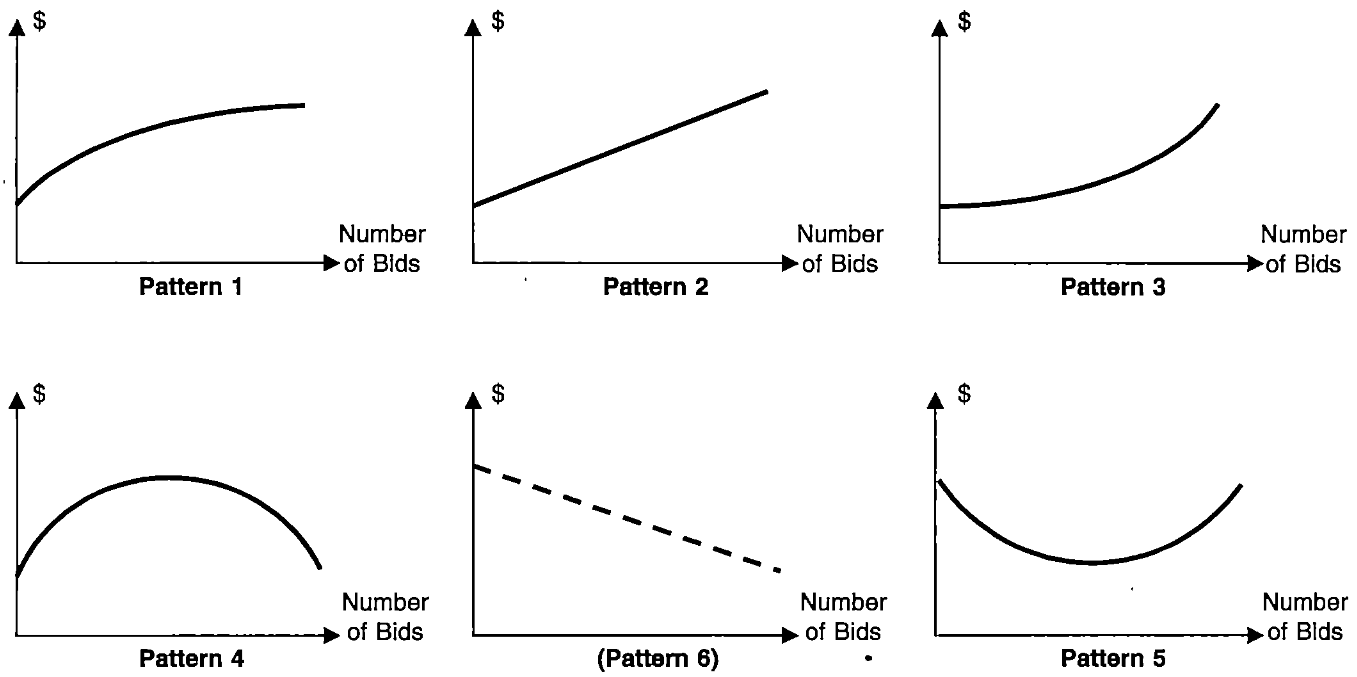
Pattern 4 reflects bids whose values follow a full inverted U shape. A sequence follows Pattern 4 if the

TABLE 3
Description of Data from Low-Cost Airline

Distribution of Bidding Sequences by Bid Length		Distribution of Consumers by Destinations	
Bid Length	**Number of Sequences**	**Destinations Bid**[a]	**Number of Consumers**
1	922	1	1937
2	981	2	377
3	600	3	110
4	322	4	22
5	227	5	—
6	156	6	1
>6	390	7	—
All	3598	All	2389

[a]Destinations for which a consumer bids.

FIGURE 2
Idealized Patterns of Quadratic Function



parameter of the quadratic term is negative ($\alpha_{1k} < 0$) and the value of the last bid is less than or equal to the value of the previous one ($B_{k,I\max} \le B_{k,I\max-1}$); that is, the curve includes the right, decreasing part of the inverted U shape. The parameter of the linear term (α_{2k}) can be positive, zero, or negative. Appendix C shows that these conditions are met when

$$(3) \qquad 0 > \alpha_{1k} \le \frac{-\alpha_{2k}}{2 \times I_{k,\max} - 1}.$$

Pattern 3 reflects bids whose values follow an increasing curve at an increasing rate (i.e., exponential increase or only the increasing part of a U-shaped curve). A sequence follows Pattern 3 if the quadratic term has a parameter of positive value ($\alpha_{1k} > 0$) and the value of the second bid of the sequence is greater than the value of the very first one ($B_{k,2} > B_{k,1}$). The parameter of the linear term (α_{2k}) can be positive, zero, or negative. Appendix C shows that these conditions are met when

$$(4) \qquad 0 < \alpha_{1k} > \frac{-\alpha_{2k}}{3}.$$

Pattern 5 reflects bids whose values follow a U-shaped curve. A sequence follows Pattern 5 if the parameter of the linear term is negative ($\alpha_{2k} < 0$), the parameter of the quadratic term is positive ($\alpha_{1k} > 0$), and the value of the second bid of the sequence is less than or equal to the value of the first one ($B_{k,2} \le B_{k,1}$); that is, the curve includes the left, decreasing part of a U-shaped curve. Appendix C shows that these conditions are met when

$$(5) \qquad 0 < \alpha_{1k} \le \frac{-\alpha_{2k}}{3}.$$

Table 4 summarizes these criteria for classification of the patterns.

Pattern 1 is consistent with strongly rational behavior (i.e., a monotonically increasing pattern with decreasing increments, as expected by the normative predictions). Pattern 3 is weakly rational (i.e., a monotonically increasing pattern with increasing increments). Pattern 2 is rational if bidding costs are zero, a condition that is unlikely to be met in real markets. Patterns 4 and 5 are clearly irrational, showing both increasing and decreasing trends in the same sequence. Similarly, Pattern 6 is clearly irrational, showing linearly declining bids.

We do not have sufficient degrees of freedom to test the significance of each of the estimated parameters in Equation 1. However, we can test the validity of our classification in two ways: First, we test the extent to which people within a group vary from the mean of that group. We do this by testing the extent to which the group means for each of the estimated parameters are statistically different from zero, given the alternate hypothesis that they take on the values we specify in Table 4. In the cases that the mean itself should be zero, as we specify in Table 4, we expect not to reject the null hypothesis. Second, we test whether the means between groups are different from each other.

Results. Table 5, Panel A, displays the estimation results of fitting the quadratic function to individual bidding sequences at the name-your-own-price retailer. On average, our model explains 89% of the variance. The results of our two tests for these data are as follows: First, in Patterns 1, 3, 4, and 5, group means for the parameters of the quadratic term (α_{1k}) and the linear term (α_{2k}) are significantly different from zero and in the expected direction, which unambiguously classifies the pattern according to the criteria that Table 4 summarizes. For Pattern 2, the mean value of the parameter of the quadratic term (α_{1k}) is not different from zero, which is exactly the characteristic we expected. We did not observe Pattern 6 in our data. Second, to test differ-

TABLE 4
Criteria for Classifying Bidding Patterns

		Range of Parameter of Quadratic Term		
		$\alpha_{1k} < 0$	$\alpha_{1k} = 0$	$\alpha_{1k} > 0$
Range of Parameter of Linear Term	$\alpha_{2k} > 0$	$0 > \alpha_{1k} > \frac{-\alpha_{2k}}{2 \times I_{k,max} - 1} \wedge \alpha_{2k} > 0$ ①	$\alpha_{1k} = 0 \wedge \alpha_{2k} > 0$ ②	$0 < \alpha_{1k} > \frac{-\alpha_{2k}}{3}$ ③
	$\alpha_{2k} = 0$	④ $0 > \alpha_{1k} \le \frac{-\alpha_{2k}}{2 \times I_{k,max} - 1}$	$\alpha_{1k} = 0 \wedge \alpha_{2k} \le 0$ ⑥	
	$\alpha_{2k} < 0$			$0 < \alpha_{1k} \le \frac{-\alpha_{2k}}{3} \wedge \alpha_{2k} < 0$ ⑤

Notes: Numbers in circles show pattern numbers (see Figure 2). Shaded areas represent clearly irrational patterns (not monotonically increasing).

TABLE 5
Proportion of Bidding Patterns

Pattern	Observations	Proportion (%)	Mean Number of Bids	Mean R^2 (%)		Mean Parameter Estimates[b] α_{1k} SE (*p* Value)	Mean Parameter Estimates[b] α_{2k} SE (*p* Value)
A. Name-Your-Own-Price Retailer							
1	511	36	4.1	94		–22.88	161.02
						2.42 (.00)	12.37 (.00)
2[a]	75	5	3.8	98		.00	77.24
						.00 (1.00)	5.95 (.00)
3	325	23	4.5	94		20.35	–23.37
						1.46 (.00)	4.22 (.00)
4	317	23	4.8	81		–36.01	174.30
						3.25 (.00)	13.71 (.00)
5	187	13	5.4	78		44.21	–187.68
						5.93 (.00)	22.69 (.00)
All	1415	100	4.6	89	χ^{2}[c]	1054.72 (.00)	839.60 (.00)
B. Low-Cost Airline							
1	808	48	5.6	98		–1.34	12.30
						.08 (.00)	.43 (.00)
2[a]	206	12	4.0	99		.00	6.59
						.00 (1.00)	.38 (.00)
3	483	29	5.2	98		1.92	–1.10
						.26 (.00)	.77 (.15)
4	140	8	7.5	82		–4.12	22.50
						.64 (.00)	2.64 (.00)
5	58	3	8.1	77		4.35	–18.42
						1.03 (.00)	4.06 (.00)
All	1695	100	5.5	96	χ^{2}[c]	1333.22 (.00)	822.96 (.00)

[a] $\alpha_{1k} = 0$.

[b] First row: mean parameter for each group of similar patterns; second row: within-group standard errors; *p* value is in parentheses.

[c] Chi-square of Kruskal–Wallis test for parameter differences between groups of similar patterns, *p* value is in parentheses.

ences of the means between groups, we used a nonparametric test (a Kruskal–Wallis test) because standard deviations differ between the groups. The test yields significant differences of α_{1k} and α_{2k} between groups (for bivariate tests of differences of means between groups, see Appendix D). Only 36% of all bidding sequences fit Pattern 1 (i.e., strongly rational), 23% fit Pattern 3 (i.e., weakly rational), and 5% fit Pattern 2 (which would be rational if bidding costs were zero). However, Hann and Terwiesch (2003) and Spann, Skiera, and Schäfers (2004) find mean bidding costs (which they call "frictional costs") of approximately $6 and $3. Furthermore, 36% of all sequences are not monotonically increasing, which is clear evidence of irrationality. Thus, overall, at the name-your-own-price retailer, 64% of the bidding sequences range from weakly rational to clearly irrational.

Table 5, Panel B, depicts the estimation results of fitting a quadratic function to individual bidding sequences at the low-cost airline. On average, our model explains 96% of the variance. The results of our two tests for these data are as follows: First, group means for the parameters of the quadratic term (α_{1k}) and the linear term (α_{2k}) are significantly different from zero, except for the parameter α_{1k} of Pattern 2 and α_{2k} of Pattern 3. We already outlined that $\alpha_{1k} = 0$ is the characteristic of Pattern 2. In Appendix C, we outline that in case of Pattern 3, α_{2k} can take a positive value, zero, and, to a certain limit, a negative value. Thus, the insignificance of α_{2k} in this case does not reject Pattern 3. Second, the Kruskal–Wallis test yields significant differences of α_{1k} and α_{2k} between groups (for bivariate tests of differences of means between groups, see Appendix D): 48% of all bidding sequences fit Pattern 1 (strongly rational), 12% fit Pattern 2, and 29% fit Pattern 3. In addition, 11% of all sequences are not monotonically increasing. Therefore, at the low-cost airline, 52% of bidding sequences range from weakly to clearly irrational.

However, bidding at the low-cost airline shows less deviation from rational behavior than at the name-your-own-price retailer. The reason is that customers of the low-cost airline are more attuned to cost minimization and are more likely to follow the predictions of the economic model of rational bidding than are customer of the air ticket retailer.

Second-Stage Analysis

We begin by hypothesizing what factors could explain the occurrence of nonmonotonic patterns in the data. Then, we show the results of testing these hypotheses. Finally, we attempt to explain the values of the bids.

Explanation of nonmonotonic patterns. This subsection explains the factors that lead to monotonically increasing patterns. The dependent variable is the presence of a monotonically increasing bid pattern. The independent variables are the number of bids in a sequence, the number of destinations for which a consumer bids, the mean time between each consecutive bid of a bidding sequence (i.e., the mean interbid time), and the region of the destination (see Equation 6). For each independent variable, we develop a hypothesis for its effect on the dependent variable. We use a binary logit model for the analysis. Thus:

$$(6)\qquad \mathrm{PR}(\mathrm{Mon}_k) = \frac{1}{1+\exp\left[-\left(\begin{array}{l}\beta_0+\beta_1\times\mathrm{NBid}_k+\beta_2\times\mathrm{NDes}_{j,k}\\ +\beta_3\times\mathrm{ITime}_k+\beta_4\times\mathrm{Euro}_k\\ +\beta_5\times\mathrm{NorAm}_k+\beta_6\times\mathrm{RWor}_k\end{array}\right)\right]}\quad (k\in K),$$

where

$\mathrm{PR}(\mathrm{Mon}_k)$ = the probability of a monotonically increasing pattern of the kth sequence,
NBid_k = the number of bids in the kth sequence,
$\mathrm{NDes}_{j,k}$ = the number of destinations for which the jth consumer in the kth sequence bid,
ITime_k = the mean time between consecutive bids in the kth sequence,
Euro_k = a dummy variable (1 if flight destination of the kth sequence is in Europe),
NorAm_k = a dummy variable (1 if flight destination of the kth sequence is in North America),
RWor_k = a dummy variable (1 if flight destination of the kth sequence is in the rest of the world, and
K = the index set of bidding sequences.

The rationale for the independent variables is the following: Consumers who bid for more destinations have more experience with the name-your-own-price auction than those who bid for less destinations. The reason is that each destination has its own configuration of dates, times, and competitors; consumers' assumptions about supply and threshold prices; and outcomes of consumers' bidding. A final acceptance indicates how low the seller's threshold price is, and a final rejection indicates how high it might be. Thus, consumers gain rich knowledge from experience with bidding for each destination (Remus, O'Connor, and Griggs 1996). Learning from experience can compensate for limited information and further rational behavior (Arthur 1991; Conlisk 1996). Thus, experienced consumers should be less prone to behavioral limitations and more likely to behave according to the normative model (i.e., rational):

> H_1: The number of destinations for which a consumer bids has a positive effect on the probability of a monotonically increasing pattern.

After we control for experience, the number of bids consumers place within a sequence may adversely affect the rationality of their bidding for several reasons. First, if consumers must take too much information into account, their mental capacity becomes limited, and their decisions can suffer (Jacoby, Speller, and Kohn 1974; Lurie 2004). As consumers make more bids, they are more likely to become confused with their prior bids. Thus, they are less likely to make rational bids. Second, bidding at a name-your-own-price auction requires consumers to develop some belief about the retailer's threshold price and update this belief when the bid is rejected, which can be difficult (Chernev 2003). Third, purely by chance, the more bids a consumer makes, the greater is the likelihood that the consumer will either make errors or become confused.

We suspect that, together, these three factors overwhelm the small increase in experience that may come from merely increasing the number of bids in each sequence. Similarly, these factors are likely to overwhelm the small decrease in learning costs that may occur with increasing lengths of bids in a sequence. Thus:

H_2. The number of bids in a sequence has a negative effect on the probability of a monotonically increasing pattern.

In general, consumers' memory and recall of prices is low (Dickson and Sawyer 1990; Monroe and Lee 1999; Rajendran and Tellis 1994; Vanhuele and Drèze 2002). As interbid time increases, consumers are more likely to forget about their prior bids. Thus, we hypothesize that the time between consecutive bids adversely affects the rationality of consumer bidding behavior:

H_3: The mean interbid time of bids in a sequence has a negative effect on the probability of a monotonically increasing pattern.

The flight time and cost of the ticket increase with the distance between the origin and the destination of a flight. Thus, the larger the distance, the greater a consumer can save by more careful planning and more rational bidding. The greater the saving, the more the consumer will be motivated to make a well-thought-out, rational decision (Tellis and Wernerfelt 1987). Thus, the longer the distance, the more a consumer will bid rationally to obtain a good deal. Because all bids were for flights originating in Germany, we operationalized distance through dummy variables for the region of the destination (Europe, North America, and the rest of the world); a flight within Germany was the baseline category:

H_{4a}: An international destination has a positive effect on the probability of a monotonically increasing pattern.

H_{4b}: The magnitude of the effect increases (becomes more positive) with the distance of the region.

Results. Table 6 shows the estimation results of the binary Logit model for the name-your-own-price retailer (second column) and the name-your-own-price offering at the low-cost airline (third column).

In both studies, contrary to H_1, the effect of the number of destinations on the probability of a monotonically increasing pattern is not significant (and is negative). Apparently, experience does not mitigate the behavioral limitations that lead to irrational bidding behavior (e.g., Prabhu and Tellis 2000). In both studies, consistent with H_2, the number of bids in a sequence has a highly significant, negative effect on the probability of a monotonically increasing pattern. In both studies, consistent with H_3, we find that mean interbid time of a sequence (i.e., the time between consecutive bids) has a negative influence on the probability of a monotonically increasing pattern. We measure interbid time at the name-your-own-price retailer because of the availability at the daily level, whereas we have data on the level of seconds for the low-cost airline. In Study 2, consistent with H_4, an increasing flight distance leads to a higher likelihood of a monotonically increasing pattern. Thus, the magnitude of the effect (positive value of the parameter of the dummy variable) increases with the distance (the rest of the world consists of destinations in South America, Africa, Asia, and Australia). We cannot test H_4 for the second data set of the name-your-own-price offering at the low-cost airline, because all flights were within Europe and had about the same distance (between one and two hours of flight time).

Both studies yield identical results for the factors that influence deviations from normative bidding behavior, indicating robust effects. However, the proportion of such deviations is much higher at the airline ticket retailer than at the low-cost airline (see Table 5). At the retailer, consumers bid for a flight that could be served by various airlines and had no knowledge of the actual airline, which was revealed only after a successful purchase (Priceline uses a similar policy). Thus, the task of generating the bid values is more difficult because of this greater uncertainty. At the airline, consumers could bid for a flight only at this specific low-cost airline; thus, they knew the product. Furthermore, the total duration of the offering at the low-cost airline was less than four days (i.e., there was less time to forget about previous bids).

Explanation of bid values. This section analyzes the maximum bid value of a bidding sequence. The indepen-

TABLE 6
Results of Binary Logit Model (Probability of Monotonically Increasing Patterns)

Parameters[b]	Retailer (Study 1)[a]	Low-Cost Airline (Study 2)[a]
Constant	1.17 (.00)	2.85 (.00)
Number of destinations	−.06 (.34)	−.16 (.11)
Number of bids	−.12 (.00)	−.09 (.00)
Mean interbid time[c]	−.03 (.00)	−.08 (.01)
DV_Europe[d]	.30 (.06)	
DV_North America[d]	.49 (.01)	
DV_rest of world[d]	.66 (.00)	
Likelihood ratio test: χ^2	123.72 (.00)	46.77 (.000)
Classification goodness (%)	69.4	88.4
Number of observations	1415	1695

[a]*p* value is in parentheses.
[b]Dependent variable. Prob(Y = 1) is the probability of monotonically increasing pattern.
[c]Mean time difference in days (hours) between consecutive bids of a sequence at retailer (low-cost airline).
[d]Baseline category. Germany

dent variables are the same as in the analysis of nonmonotonic patterns. We use multiple regression to analyze the dependent variable:

$$(7)\quad \text{MaxBidV}_k = \beta_0 + \beta_1 \times \text{NBid}_k + \beta_2 \times \text{NDes}_{j,k} + \beta_3 \times \text{ITime}_k + \beta_4 \times \text{Euro}_k + \beta_5 \times \text{NorAm}_k + \beta_6 \times \text{RWor}_k + \mu_k \quad (k \in K),$$

where MaxBidV_k is the maximum bid value of the kth sequence and μ_k is the residual of the kth sequence.

Table 7 shows the estimation results of the regression analysis for the maximum bid value of a bidding sequence for the name-your-own-price retailer (second column) and for the name-your-own-price offering at the low-cost airline (third column). In both studies, the maximum bid value decreases with the number of destinations. Apparently, inexperienced consumers may overbid. Conversely, experienced consumers who bid for more than one destination are more likely to bid lower. Thus, experience does not lead to a higher likelihood of rational bidding behavior but reduces the consumer's maximum bid value.

In both studies, the maximum bid value of a sequence increases with the number of bids. Thus, even if consumers become confused with an increasing number of bids, they raise their average bid by the same amount and overbid. Interbid time has no significant effect on the maximum bid value in both studies. Apparently, the waiting time between consecutive bids influences the rationality of the bidding pattern but does not have an effect on the maximum bid value. The distance of a flight increases maximum bid value because consumers correctly anticipate that costs (especially fuel costs) increase with distance and that higher costs lead to a higher threshold price for a flight. Again, we cannot analyze this effect for the data for the low-cost airline because of similar distances of all destinations.

We found that the same factors explain maximum bid value in both studies. This indicates the stability of our results for name-your-own-price auctions because both studies differ with respect to currency, time of the study, duration of data gathering, and offering (multiple airlines versus one low-cost airline).

Discussion

The popularity of the Internet is supposed to have helped consumers make better decisions that are more consistent with rational economic theory than was possible before the Internet. We analyze bidding sequences for airline tickets at a name-your-own-price retailer and at a low-cost airline. We find that a majority of bidding sequences are not consistent with the predictions of an economic model of a rational, price-minimizing consumer. This finding indicates that the Internet does not eliminate or lower consumers' irrational decisions as many experts expected or hoped. Further analysis reveals that the deviations from rationality are mostly consistent with a priori expectations. In particular, consumers who place many bids with rather long interbid times are more likely to bid irrationally, which can be attributed to forgetting. Consumers are more likely to bid rationally for larger distances, which can be explained by higher involvement for or savings from such flights. Consumers who have more experience have lower bids on average, probably because of the knowledge they have gained. Our results have some important implications for managers, consumers, and researchers.

Implications for Managers

Price-discovery mechanisms, such as name-your-own-price auctions, are facilitated by the low transaction costs of trade on the Internet. Such mechanisms promise sellers two important benefits: First, prices based on the bidding process can be consumer specific and allow for better extraction of consumer surplus through this segmentation. Second, price discovery mechanisms can reduce efforts by third parties that post lowest market prices and thus thwart firms' price discrimination of consumers.

Our results indicate that despite the convenience of the Internet to consumers, most consumers still do not make strictly rational decisions. This finding suggests that firms

TABLE 7
Regression Results for Maximum Bid Value

Parameters	Retailer (Study 1)[a, d]	Low-Cost Airline (Study 2)[a, e]
Constant	265.93 (.000)	45.20 (.000)
Number of destinations	–8.80 (.098)	–3.73 (.000)
Number of bids	3.45 (.083)	1.57 (.000)
Mean interbid time[b]	–.08 (.808)	.55 (.172)
DV_Europe[c]	121.73 (.000)	
DV_North America[c]	632.38 (.000)	
DV_rest of world[c]	901.61 (.000)	
R^2	.694	.059
F-value	531.22 (.000)	35.32 (.000)
Number of observations	1415	1695

[a] *p* value is in parentheses

[b] Mean time difference in days (hours) between consecutive bids of a sequence at retailer (low-cost airline).

[c] Baseline category: Germany.

[d] Bid value in Deutschmark (DM) (DM1 = approximately $.5).

[e] Bid value in euro (€1 = approximately $1.2)

can further segment consumers by their bidding patterns and exploit their irrationality for profit. For this purpose, firms would need to modify the design of such auctions or develop decision rules that alter acceptance of bids on the basis of the bidding pattern. For example, if enough seats are projected to be available, firms might accept bids early from those consumers who show declining bids, as long as the price is over the firms' threshold price. Alternatively, they could hold off accepting bids from consumers who seem to be in a panic mode and show increasing bids at an increasing rate. However, such a strategy might not be successful if anticipated by consumers who then try to conceal their true characteristics. In general, firms should not rely solely on models of behavior that assume strict rationality, because such models may not reflect reality and may recommend suboptimal pricing strategies.

Implications for Consumers

The Internet may allow for more informed shopping and more efficient price discovery mechanisms than before. In particular, name-your-own-price auctions promise consumers the option of obtaining a price that is below their reservation price but close to a seller's marginal cost. However, many consumers may not fully exploit these price discovery mechanisms to obtain as low a price as they hoped. This state of affairs may exist because these mechanisms are new, because consumers do not fully understand how they run, or because consumers are not experienced enough to exploit them fully. For example, we find that inexperienced consumers tend to overbid in name-your-own-price auctions, thus offering higher prices to sellers that use such mechanisms. Consumer advocate groups could issue warnings of the irrational bidding patterns that consumers are prone to make in such auctions. They could also develop advisories that educate consumers on how best to use such systems to retain as much consumer surplus as possible given supply constraints and their own urgency to travel.

Implications for Researchers

The assumption of rationality in economic models has been debated for decades. Behavioral economists, such as Simon (1955), Kahneman and Tversky (1979), and Thaler (1985), have not only questioned the validity of this assumption but also proposed alternate theories that do not depend on the assumption of strict rationality adopted in traditional economic models. Our results provide doubts about the assumption of strict rationality even in the context of the Internet, which is supposed to be a boon to consumers. As such, models that incorporate more realistic assumptions of consumer decision making are likely to have enhanced relevance and predictive accuracy.

Furthermore, such enhanced models may enable online auctioneers to improve revenue predictions or to extract consumer surplus more effectively through higher threshold prices, given predictions on the maximum bid values consumers submit. In addition, these models can allow the development of price discovery mechanisms that support consumer decision making and attain higher efficiency.

Although our results are robust across our two empirical studies of consumer bidding behavior, they are limited by geography and context. Replications across other markets and contexts would be useful to explore nonnormative behavior on the Internet further.

In addition to econometric analyses of behavioral data, this phenomenon lends itself to fruitful experimental research, in the laboratory and on the Internet. Laboratory experiments allow for the testing of specific theories as to why consumers behave in irrational ways. The Internet enables field experiments easily. Such field experiments can explore what ecological contexts promote or mitigate consumers' irrational behavior and how organizations can avoid or exploit such behavior.

Finally, our theoretical model does not take into account gaming strategies by firms that discover consumers' bidding patterns or consumers who expect firms to do so and thus hide their patterns. Such models make for interesting game theoretic solutions and might be worth pursuing.

Appendix A
Consumers' Optimal Bid Prices

The consumer optimizes the expected consumer surplus of the bid over the bid amount, solving Equation A1. A jth consumer's belief (i.e., expectation) about the seller's threshold price p is assumed to follow a uniform distribution on the interval $[\underline{p}_j, \overline{p}_j]$ with $\overline{p}_j \geq r_j$, where r_j is the consumer's reservation price (Ding et al. 2005; Hann and Terwiesch 2003; Spann, Skiera, and Schäfers 2004; Stigler 1961).

$$\text{(A1)} \quad \max_{b_{j,1},\;,b_{j,n_j}} ECS_{j,1} = (r_j - b_{j,1}) \times \text{Prob}(b_{j,1} \geq p) + ECS_{j,2} \times \text{Prob}(b_{j,1} < p) - c_{j,1}$$

$$= (r_j - b_{j,1}) \times \frac{b_{j,1} - \underline{p}_j}{\overline{p}_j - \underline{p}_j} + ECS_{j,2} \times \frac{\overline{p}_j - b_{j,1}}{\overline{p}_j - \underline{p}_j} - c_{j,1}$$

$$\text{subject to } ECS_{j,i} \geq 0,\ b_{j,i} \leq r_j \ \forall i \leq n_j.$$

The jth consumer's expected consumer surplus $ECS_{j,1}$ of the first bid $b_{j,1}$ has two components in this model (Equation A1). The first component represents the expected consumer surplus in case of a successful bid, which is weighted with the probability that the first bid is successful. Thus, the expected consumer surplus accounts for the bidding costs $c_{j,i}$, which accrue by submitting the ith bid. The second component illustrates the expected consumer surplus $ECS_{j,2}$ of a second bid, which is weighted with the probability that the first bid is unsuccessful. As such, $ECS_{j,2}$ consists of the expected consumer surplus of a second bid and further bids beyond the second, if the previous bids are not successful and if it is still beneficial for the bidder to make these additional bids. Thus, the consumer surplus of further bids beyond the ith bid is recursively included in the formula for the consumer surplus of the ith bid.

The consumer simultaneously determines the optimal number of bids and the optimal bid prices by solving this

TABLE A1
Bid Values of Optimal Bidding Strategies for One to Six Bids

	One Bid	**Two Bids**
First bid	$b^*_{j,1} = \frac{1}{2}\underline{p}_j + \frac{1}{2}r_j$	$b^*_{j,1} = \frac{2}{3}\underline{p}_j + \frac{1}{3}r_j + \frac{2}{3}c_{j,2}$
Second bid		$b^*_{j,2} = \frac{1}{3}\underline{p}_j + \frac{2}{3}r_j + \frac{1}{3}c_{j,2}$
	Three Bids	**Four Bids**
First bid	$b^*_{j,1} = \frac{3}{4}\underline{p}_j + \frac{1}{4}r_j + \frac{3}{4}c_{j,2} + \frac{2}{4}c_{j,3}$	$b^*_{j,1} = \frac{4}{5}\underline{p}_j + \frac{1}{5}r_j + \frac{4}{5}c_{j,2} + \frac{3}{5}c_{j,3} + \frac{2}{5}c_{j,4}$
Second bid	$b^*_{j,2} = \frac{2}{4}\underline{p}_j + \frac{2}{4}r_j + \frac{2}{4}c_{j,2} + \frac{4}{4}c_{j,3}$	$b^*_{j,2} = \frac{3}{5}\underline{p}_j + \frac{2}{5}r_j + \frac{3}{5}c_{j,2} + \frac{6}{5}c_{j,3} + \frac{4}{5}c_{j,4}$
Third bid	$b^*_{j,3} = \frac{1}{4}\underline{p}_j + \frac{3}{4}r_j + \frac{1}{4}c_{j,2} + \frac{2}{4}c_{j,3}$	$b^*_{j,3} = \frac{2}{5}\underline{p}_j + \frac{3}{5}r_j + \frac{2}{5}c_{j,2} + \frac{4}{5}c_{j,3} + \frac{6}{5}c_{j,4}$
Fourth bid		$b^*_{j,4} = \frac{1}{5}\underline{p}_j + \frac{4}{5}r_j + \frac{1}{5}c_{j,2} + \frac{2}{5}c_{j,3} + \frac{3}{5}c_{j,4}$
	Five Bids	**Six Bids**
First bid	$b^*_{j,1} = \frac{5}{6}\underline{p}_j + \frac{1}{6}r_j$ $+ \frac{5}{6}c_{j,2} + \frac{4}{6}c_{j,3} + \frac{3}{6}c_{j,4} + \frac{2}{6}c_{j,5}$	$b^*_{j,1} = \frac{6}{7}\underline{p}_j + \frac{1}{7}r_j + \frac{6}{7}c_{j,2} + \frac{5}{7}c_{j,3}$ $+ \frac{4}{7}c_{j,4} + \frac{3}{7}c_{j,5} + \frac{2}{7}c_{j,6}$
Second bid	$b^*_{j,2} = \frac{4}{6}\underline{p}_j + \frac{2}{6}r_j$ $+ \frac{4}{6}c_{j,2} + \frac{8}{6}c_{j,3} + \frac{6}{6}c_{j,4} + \frac{4}{6}c_{j,5}$	$b^*_{j,2} = \frac{5}{7}\underline{p}_j + \frac{2}{7}r_j + \frac{5}{7}c_{j,2} + \frac{10}{7}c_{j,3}$ $+ \frac{8}{7}c_{j,4} + \frac{6}{7}c_{j,5} + \frac{4}{7}c_{j,6}$
Third bid	$b^*_{j,3} = \frac{3}{6}\underline{p}_j + \frac{3}{6}r_j$ $+ \frac{3}{6}c_{j,2} + \frac{6}{6}c_{j,3} + \frac{9}{6}c_{j,4} + \frac{6}{6}c_{j,5}$	$b^*_{j,3} = \frac{4}{7}\underline{p}_j + \frac{3}{7}r_j + \frac{4}{7}c_{j,2} + \frac{8}{7}c_{j,3}$ $+ \frac{12}{7}c_{j,4} + \frac{9}{7}c_{j,5} + \frac{6}{7}c_{j,6}$
Fourth bid	$b^*_{j,4} = \frac{2}{6}\underline{p}_j + \frac{4}{6}r_j$ $+ \frac{2}{6}c_{j,2} + \frac{4}{6}c_{j,3} + \frac{6}{6}c_{j,4} + \frac{8}{6}c_{j,5}$	$b^*_{j,4} = \frac{3}{7}\underline{p}_j + \frac{4}{7}r_j + \frac{3}{7}c_{j,2} + \frac{6}{7}c_{j,3}$ $+ \frac{9}{7}c_{j,4} + \frac{12}{7}c_{j,5} + \frac{8}{7}c_{j,6}$
Fifth bid	$b^*_{j,5} = \frac{1}{6}\underline{p}_j + \frac{5}{6}r_j$ $+ \frac{1}{6}c_{j,2} + \frac{2}{6}c_{j,3} + \frac{3}{6}c_{j,4} + \frac{4}{6}c_{j,5}$	$b^*_{j,5} = \frac{2}{7}\underline{p}_j + \frac{5}{7}r_j + \frac{2}{7}c_{j,2} + \frac{4}{7}c_{j,3}$ $+ \frac{6}{7}c_{j,4} + \frac{8}{7}c_{j,5} + \frac{10}{7}c_{j,6}$
Sixth bid		$b^*_{j,6} = \frac{1}{7}\underline{p}_j + \frac{6}{7}r_j + \frac{1}{7}c_{j,2} + \frac{2}{7}c_{j,3}$ $+ \frac{3}{7}c_{j,4} + \frac{4}{7}c_{j,5} + \frac{5}{7}c_{j,6}$

optimization problem. The optimal bidding strategy is to submit a further bid, if the previous one is declined, as long as the expected consumer surplus of this bid is not negative. Thus, the optimal number of bids n_j is determined. The optimal bidding strategy can be determined by the algorithm that Spann, Skiera, and Schäfers (2004) propose.

We can state the optimal bid values for different optimal numbers of bids n_j according to the following general formula (for bid values of optimal bidding strategies for one to six bids, see Table A1):

$$\text{(A2)} \qquad b^*_{j,i|n_j} = \frac{n_j+1-i}{n_j+1} \times \underline{p}_j + \frac{i}{n_j+1} \times r_j + \sum_{\substack{l=2 \\ \text{for } i \geq 2}}^{i} \frac{(n_j+1-i) \times (l-1)}{n_j+1} \times c_{j,l} + \sum_{\substack{m=i+1 \\ \text{for } i < n_j}}^{n_j} \frac{(n_j+2-m) \times i}{n_j+1} \times c_{j,m}.$$

Because rational consumers determine *ex ante* their optimal bidding behavior, bidding costs influence their *ex ante* determination of bidding behavior and thus are not sunk.

Appendix B
Proof of Norms for Bidding Patterns

P_1: Any two consecutive bids $b^*_{j,i-1|n_j}$ and $b^*_{j,i|n_j}$ have positive increments $\Delta b^*_{j,i|n_j} = b^*_{j,i|n_j} - b^*_{j,i-1|n_j} > 0$.

Proof. Assuming constant bidding costs $c_j = c_{j,i}$ for each ith bid, we can state the general rule for the ith increment, which we derive from Equation A2:

$$\text{(B1)} \quad \Delta b^*_{j,i|n_j} = b^*_{j,i|n_j} - b^*_{j,i-1|n_j} = \frac{1}{n_j+1} \times (r_j - \underline{p}_j) + \frac{c_j}{n_j+1} \times \left[\frac{(n_j-1) \times n_j}{2} - 2 - (n_j+1) \times (i-2)\right]$$
$$= \frac{1}{n_j+1} \times (r_j - \underline{p}_j) + c_j \times \left[\frac{(n_j-1) \times n_j - 4}{2 \times (n_j+1)} - (i-2)\right].$$

Because Equation B1 decreases in the bid number i (for positive costs, which we develop a proof for subsequently), the increment value is lowest for the optimal bid number $i = n_j$. Thus, increments are positive for all $i \leq n_j$ if (inserting $i = n_j$ into Equation B1):

$$\text{(B2)} \quad \Delta b^*_{j,i=n|n_j} = \frac{1}{n_j+1} \times (r_j - \underline{p}_j) + \frac{c_j}{n_j+1} \times \left[\frac{(n_j-1) \times n_j}{2} - 2 - (n_j+1) \times (n_j-2)\right] > 0$$
$$= \frac{1}{n_j+1} \times \left[(r_j - \underline{p}_j) + c_j \times \frac{n_j - n_j^2}{2}\right]$$
$$> 0 \Leftrightarrow r_j - \underline{p}_j > c_j \times \frac{n_j^2 - n_j}{2}.$$

Thus, for bidding costs of zero, increments are positive for $r_j > \underline{p}_j$. The value of Equation B2 could become negative for high values of c_j and n_j. However, surplus-maximizing consumers would never place such a bid with a negative increment relative to the previous bid, because it would lead to a negative expected consumer surplus of this specific bid:

$$\text{(B3)} \qquad ECS_{j,i} = (r_j - b_{j,i}) \times \frac{b_{j,i} - b_{j,i-1}}{\overline{p}_j - b_{j,i-1}} - c_j < 0$$
$$\text{for } b_{j,i} - b_{j,i-1} < 0 \text{ and } \overline{p}_j > b_{j,i-1}.$$

The second norm for the rational bidding pattern is given in P_1. If any two consecutive bids have positive increments, the first bid a consumer submits is the lowest, reflecting the first norm.

P_2: The increments between consecutive bids decrease by the amount of the bidding costs. Thus, the normative bidding pattern has decreasing increments for positive bidding costs.

$$\text{(B4)} \qquad \Delta b^*_{j,i|n_j} - \Delta b^*_{j,i-1|n_j} = -c_j < 0 \text{ for } c_j > 0.$$

Proof. If we assume constant bidding costs $c_j = c_{j,i}$ for each ith bid,

$$\text{(B5)} \quad \Delta b^*_{j,i|n_j} - \Delta b^*_{j,i-1|n_j} = \left(b^*_{j,i|n_j} - b^*_{j,i-1|n_j}\right) - \left(b^*_{j,i-1|n_j} - b^*_{j,i-2|n_j}\right)$$
$$= \frac{1}{n_j+1} \times (r_j - \underline{p}_j) + c_j \times \left[\frac{(n_j-1) \times n_j - 4}{2 \times (n_j+1)} - (i-2)\right] - \left(\frac{1}{n_j+1} \times (r_j - \underline{p}_j) + c_j \times \left\{\frac{(n_j-1) \times n_j - 4}{2 \times (n_j+1)} - [(i-1)-2]\right\}\right)$$
$$= c_j[-(i-2) + (i-3)]$$
$$= c_j(-i+2+i-3) = c_j(-1) = -c_j.$$

or

$$\text{(B6)} \quad \cdot \qquad \frac{\partial \Delta b^*_{j,i|n_j}}{\partial i} = -c_j.$$

Thus, for positive and constant bidding costs, increments are decreasing on the basis of P_2, which is the third norm for the rational bidding pattern. For bidding costs of zero, increments are constant (and positive).

Appendix C
Proof for Criteria for Classification of Four Complex Bidding Patterns

Pattern 1

Pattern 1 reflects increasing bids at a decreasing rate; the last bid is still the highest in the sequence (i.e., only the upward portion of an inverted U shape). A sequence follows Pattern 1 if the parameter of the linear term is positive ($\alpha_{2k} > 0$), the parameter of the quadratic term is negative ($\alpha_{1k} < 0$), and the value of the last bid of the sequence is the highest ($B_{k,I\max} > B_{k,I\max-1}$). Inserting Equation 1 for bid values, this condition can be transformed as follows:

$$
\begin{aligned}
\text{(C1)}\quad & \alpha_{1k}\times(I_{k,\max})^2+\alpha_{2k}\times I_{k,\max}+\alpha_{3k}>\alpha_{1k}\times(I_{k,\max}-1)^2\\
& +\alpha_{2k}\times(I_{k,\max}-1)+\alpha_{3k}\\
& \Leftrightarrow \alpha_{1k}\times(I_{k,\max})^2+\alpha_{2k}\times I_{k,\max}\\
& \quad >\alpha_{1k}\times[(I_{k,\max})^2-2\times I_{k,\max}+1]+\alpha_{2k}\times(I_{k,\max}-1)\\
& \Leftrightarrow 0>\alpha_{1k}\times(-2\times I_{k,\max}+1)-\alpha_{2k}\\
& \Leftrightarrow 0>\alpha_{1k}>\frac{-\alpha_{2k}}{2\times I_{k,\max}-1}.
\end{aligned}
$$

Pattern 4

Pattern 4 reflects the full inverted U shape. A sequence follows Pattern 4 if the parameter of the quadratic term is negative ($\alpha_{1k} < 0$) and the value of the last bid is less than or equal to the value of the previous one ($B_{k,I\max} \leq B_{k,I\max-1}$); that is, the right, decreasing part of the inverted U shape is included. The parameter of the linear term (α_{2k}) can be positive, zero, or negative. Inserting Equation 1 for bid values, this condition can be transformed analogously to Pattern 1:

$$\text{(C2)}\qquad 0>\alpha_{1k}\leq\frac{-\alpha_{2k}}{2\times I_{k,\max}-1}.$$

Pattern 3

Pattern 3 reflects an increasing curve at an increasing rate (i.e., exponential increase or only the increasing part of a U-shaped curve). A sequence follows Pattern 3 if the quadratic term has a parameter of positive value ($\alpha_{1k} > 0$) and the value of the second bid of the sequence is greater than the value of the first one ($B_{k,2} > B_{k,1}$). The parameter of the linear term (α_{2k}) can be positive, zero, or negative. Inserting Equation 1 for bid values, this condition can be transformed as follows:

$$
\begin{aligned}
\text{(C3)}\quad & \alpha_{1k}\times 2^2+\alpha_{2k}\times 2+\alpha_{3k}>\alpha_{1k}\times 1^2+\alpha_{2k}\times 1+\alpha_{3k}\\
& \Leftrightarrow \alpha_{1k}\times 4+\alpha_{2k}\times 2>\alpha_{1k}\times 1+\alpha_{2k}\times 1\\
& \Leftrightarrow \alpha_{1k}\times 3+\alpha_{2k}\times 1>0\\
& \Leftrightarrow 0<\alpha_{1k}>\frac{-\alpha_{2k}}{3}.
\end{aligned}
$$

Pattern 5

Pattern 5 reflects a U-shaped curve. A sequence follows Pattern 5 if the parameter of the linear term is negative ($\alpha_{2k} < 0$), the parameter of the quadratic term is positive ($\alpha_{1k} > 0$), and the value of the second bid of the sequence is less than or equal to the value of the first one ($B_{k,2} \leq B_{k,1}$), that is, the left, decreasing part of a U-shaped curve. Inserting Equation 1 for bid values, analogously to Pattern 3, the conditions for Pattern 5 reduce to the following:

$$\text{(C4)}\qquad 0<\alpha_{1k}\leq\frac{-\alpha_{2k}}{3}.$$

Appendix D
Bivariate Tests of Differences of Means Between Groups

Table D1 displays the results for bivariate t-tests (accounting for different distributions within groups) of differences of means between groups for α_{1k} and α_{2k}. Thus, each cell in

TABLE D1
Bivariate Tests of Differences of Means Between Groups

	Parameter α_{1k}[a]				Parameter α_{2k}[a]			
Pattern	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
A. Name-Your-Own-Price Retailer								
2	–9.44 (.00)				6.10 (.00)			
3	–15.27 (.00)	–13.91 (.00)			14.11 (.00)	13.79 (.00)		
4	3.24 (.00)	11.10 (.00)	15.83 (.00)		–.71 (.47)	–6.49 (.00)	–13.78 (.00)	
5	–10.48 (.00)	–7.46 (.00)	–3.91 (.00)	–11.87 (.00)	13.49 (.00)	11.29 (.00)	7.12 (.00)	13.65 (.00)
B. Low-Cost Airline								
2	–17.81 (.00)				9.96 (.00)			
3	–12.12 (.00)	–7.44 (.00)			15.24 (.00)	8.95 (.00)		
4	4.29 (.00)	6.40 (.00)	8.71 (.00)		–3.81 (.00)	–5.97 (.00)	–8.59 (.00)	
5	–5.49 (.00)	–4.20 (.00)	–2.28 (.03)	–6.95 (.00)	–7.53 (.00)	6.14 (.00)	4.20 (.00)	8.46 (.00)

[a] Independent sample bivariate t-test (accounting for different distributions within groups) of differences of means between groups; *p* value is in parentheses.

Table D1 gives the t-value (and p value) for a bivariate test of differences of means between groups for parameter α_{1k} (left matrix) and parameter α_{2k} (right matrix) for both studies (Table D1, Panel A; Table D1, Panel B). Table D1 shows that all bivariate group means are significantly different from one another except for the means of parameter α_{2k} between Pattern 1 and Pattern 4 in Study 1 (name-your-own-price retailer).

REFERENCES

Arthur, Brian W. (1991), "Designing Economic Agents That Act Like Human Agents: A Behavioral Approach to Bounded Rationality," *American Economic Review*, 81 (2), 353–59.

Bakos, J. Yannis (1997), "Reducing Buyer Search Costs: Implications for Electronic Marketplaces," *Management Science*, 43 (12), 1676–92.

Brynjolfsson, Erik and Michael D. Smith (2001), "The Great Equalizer? Consumer Choice Behavior at Internet Shopbots," working paper, Massachusetts Institute of Technology.

Chandy, Rajesh, Gerard J. Tellis, Debbie MacInnis, and Pattana Thaivanich (2001), "What to Say When: Advertising Appeals in Evolving Markets," *Journal of Marketing Research*, 38 (November), 399–414.

Chernev, Alexander (2003), "Reverse Pricing and Online Price Elicitation Strategies in Consumer Choice," *Journal of Consumer Psychology*, 13 (1–2), 51–62.

Conlisk, John (1996), "Why Bounded Rationality?" *Journal of Economic Literature*, 34 (2), 669–700.

Dickson, Peter R. and Alan G. Sawyer (1990), "The Price Knowledge and Search of Supermarket Shoppers," *Journal of Marketing*, 54 (July), 42–53.

Ding, Min, Jehoshua Eliashberg, Joel Huber, and Ritesh Saini (2005), "Emotional Bidders: An Analytical and Experimental Examination of Consumers' Behavior in Priceline-Like Reverse Auctions," *Management Science*, 51 (3), 352–64.

Fay, Scott (2004), "Partial Repeat Bidding in the Name-Your-Own-Price Channel," *Marketing Science*, 23 (3), 407–418.

Geoffrion, Arthur M. and Ramayya Krishnan (2003), "E-Business and Management Science: Mutual Impacts (Part 1 of 2)," *Management Science*, 49 (10), 1275–86.

Hann, Il-Horn and Christian Terwiesch (2003), "Measuring the Frictional Costs of Online Transactions: The Case of a Name-Your-Own-Price Channel," *Management Science*, 49 (11), 1563–79.

Jacoby, Jacob, Donald E Speller, and Carol A. Kohn (1974), "Brand Choice Behavior as a Function of Information Overload," *Journal of Marketing Research*, 11 (February), 63–69.

Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47 (2), 263–91.

Lurie, Nicholas H. (2004), "Decision Making in Information-Rich Environments: The Role of Information Structure," *Journal of Consumer Research*, 30 (4), 473–86.

Monroe, Kent B. and Angela Y. Lee (1999), "Remembering Versus Knowing: Issues in Buyers' Processing of Price Information," *Journal of the Academy of Marketing Science*, 27 (2), 207–225.

Prabhu, Jaideep and Gerard J. Tellis (2000), "Do Consumers Ever Learn? Findings from an Experimental Market," *Journal of Behavioral Decision Making*, 13 (1), 19–34.

Rajendran, K.N. and Gerard J. Tellis (1994), "Is Reference Price Based on Context or Experience? An Analysis of Consumers' Brand Choices," *Journal of Marketing*, 58 (April), 10–22.

Remus, William, Marcus O'Connor, and Kenneth Griggs (1996), "Does Feedback Improve the Accuracy of Recurrent Judgmental Forecasts?" *Organizational Behavior and Human Decision Processes*, 66 (1), 22–30.

Shugan, Stephen (1980), "The Cost of Thinking," *Journal of Consumer Research*, 7 (2), 99–111.

Simon, Herbert A. (1955), "A Behavioral Model of Rational Choice," *The Quarterly Journal of Economics*, 69 (1), 99–118.

Spann, Martin, Bernd Skiera, and Björn Schäfers (2004), "Measuring Individual Frictional Costs and Willingness-to-Pay via Name-Your-Own-Price Mechanisms," *Journal of Interactive Marketing*, 18 (4), 22–36.

Stigler, George J. (1961), "The Economics of Information," *Journal of the Political Economy*, 69 (3), 213–25.

Tellis, Gerard J (1986), "Beyond the Many Faces of Price: An Integration of Pricing Strategies," *Journal of Marketing*, 50 (October), 146–60.

——— and Birger Wernerfelt (1987), "Competitive Price and Quality Under Asymmetric Information," *Marketing Science*, 6 (3), 240–53

Thaler, Richard H. (1985), "Mental Accounting and Consumer Choice," *Marketing Science*, 4 (3), 199–214.

Vanhuele, Marc and Xavier Drèze (2002), "Measuring the Price Knowledge Shoppers Bring to the Store," *Journal of Marketing*, 66 (October), 72–85.

Anthony D. Cox, Dena Cox, & Gregory Zimet

Understanding Consumer Responses to Product Risk Information

Two experiments examine how message framing moderates consumer responses to product risk information. The findings suggest that contrary to an influential theory, consumers exposed to loss-framed messages exhibit a general aversion to product risk involving both short-term adverse effects and more permanent harm. In contrast, consumers exposed to gain-framed messages differentiate among different types of product risk. They essentially ignore temporary product risks but give considerable decisional weight to risks of permanent harm. This article discusses the implications of these findings for those who design and regulate promotional messages that contain product risk disclosures.

With increasing frequency, consumers are asked to evaluate product risk information (i.e., information that product use may lead to negative consequences). For example, advertisements for tobacco products have long been required to disclose the health hazards associated with their use, and promotional materials for financial investments often contain caveats about potential financial losses (Fitch 1997; Fox et al. 1998). However, the most dramatic growth in advertised product risk disclosures has come from a relatively new category of advertising: direct-to-consumer (DTC) advertisements for complex medical products. Over the past 15 years, there has been an explosion of such advertising; total expenditures on DTC prescription drug advertising has grown more than 200-fold, from $12 million in 1989 to more than $3 billion in 2003 (Calfee 2002; Thomaselli 2004). Most such advertising is required to contain information on product risks. However, as Cohen (2002, p. 172) notes, there is "continuing concern that consumers may not adequately comprehend the benefits and risks of using such highly technical products."

There are two kinds of product risk information that consumers may consider when evaluating such products. First, there are product risk disclosures (i.e., explicit statements of the potential adverse outcomes resulting from product use). For example, Food and Drug Administration rules require that any DTC advertisement that mentions both the name of a prescription drug (e.g., Celebrex) and its function (e.g., treatment of osteoarthritis) must also contain information on product side-effect risk—that is, the possibility of unintended physical harm (ranging from temporary discomfort to permanent disability or even death) resulting from product use (for discussions of the rules governing such risk disclosures, see Calfee 2002; Menon et al. 2003b).

Second, consumers may also make inferences about the risks of specific products, even when those risks are not explicitly stated. For example, a consumer may perceive that certain product forms (e.g., injections, radiological scanning) entail inherent risks. In addition, a growing number of researchers have suggested that consumers perceive certain product functions (e.g., problem detection) as inherently riskier than others (see Meyerowitz and Chaiken 1987; Rothman and Salovey 1997).

It is important for marketers and policy makers to understand how consumers evaluate product risk information. Unfortunately, research on this topic has yielded inconsistent findings. Some studies suggest that information on product risks lowers consumers' product evaluations (e.g., Morris, Mazis, and Brinberg 1989; see also Goetzl 2001); however, others suggest that consumers frequently ignore product risk information (e.g., Menon et al. 2003a), and still others suggest that product risk information can actually enhance consumers' product evaluations (e.g., Goetzl 2000; Rodale Publications 1999).

In an effort to explain these contradictory findings, the current study proposes a potential contingency variable that may moderate the effect of product risk information on consumers' product evaluations: message framing. Any promotional message can stress either the potential gains from adopting a behavior or the potential losses from not adopting it. For example, a loss-framed advertisement promoting flu vaccinations would stress the potential health costs of not getting vaccinated (e.g., getting influenza), whereas a gain-framed advertisement would stress the protective benefits of getting vaccinated (e.g., avoiding influenza). Although these two messages might contain the same essential information, research suggests that they elicit different consumer responses (for a review of framing research, see Levin, Schneider, and Gaeth 1998). Relevant to the current study, there are theoretical reasons to believe that message framing may moderate how consumers respond to product risk.

Anthony D. Cox is Professor of Marketing, Kelley School of Business Indianapolis (e-mail: acox@iu.edu), Dena Cox is Professor of Marketing and Eli Lilly Faculty Fellow, Kelley School of Business Indianapolis (e-mail: dcox@iu.edu), and Gregory Zimet is Professor of Pediatrics and Clinical Psychology, School of Medicine (e-mail: gzimet@iupui.edu), Indiana University. A portion of this research was funded by a National Institutes of Health grant (No. R01 AI049644). The authors thank Sue Mantel for her helpful comments on a previous draft of this article.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Function/risk theory (e.g., Meyerowitz and Chaiken 1987; Rothman and Salovey 1997) draws on Kahneman and Tversky's (1979) prospect theory to suggest that negatively framed messages increase consumers' willingness to use high-risk products (particularly those with a problem detection function), whereas positively framed messages increase the attractiveness of low-risk products. In contrast, other theories (e.g., Adaval 2001; Isen 2000; Yuen and Lee 2003) seem to suggest the opposite, positing that consumers exposed to positive, upbeat stimuli should be more willing to tolerate certain types of product risks, whereas those exposed to negative stimuli should exhibit greater aversion to such risks.

This article examines these theories and discusses their (sometimes conflicting) implications for understanding how consumers evaluate product risk. Specifically, we address three central research questions:

1. How do consumers assess product risk? Specifically, how are product risk perceptions influenced by both explicit product risk disclosures and information on product function (e g., problem prevention versus problem detection).
2. How do consumers weigh perceived product risks in forming overall product evaluations?
3. How are the weights given to product risks influenced by message framing?

In the following sections, we discuss previous research and theory relevant to these questions. Then, we report the results of two experiments that examine these issues empirically. Finally, we discuss the implications of our findings and provide directions for further research.

Effects of Product Function on Perceived Product Risk

Several researchers have suggested that consumers perceive detection products as inherently risky. Meyerowitz and Chaiken (1987) first introduced the function/risk hypothesis in the context of breast self-examination (BSE). They noted (p. 89), "In the short run, BSE is a risky behavior.... [It] requires that a woman risk aversive consequences in the present (e.g., finding a lump, experiencing anxiety) in hopes of enhancing future outcomes.... Of course, in the long run, not performing BSE clearly entails risks (e.g., failing to detect a malignancy). However, because of their temporal remoteness, such longer-term considerations may be less salient than short-term considerations." More recently, Rothman and Salovey (1997) reiterated the link between product function and perceived risk, stating that detection products are considered inherently risky because "one runs the risk of receiving significant, unpleasant information" (p. 5), whereas prevention products (e.g., sunscreen, vaccines) are considered low risk because their "salient function ... is to provide a relatively certain, desirable outcome" (p. 11).

The hypothesized link between product function and perceived risk is argued persuasively in these articles and has gained widespread currency among persuasion researchers and practitioners (as we discuss subsequently). However, closer examination reveals problems with this hypothesis. First, proponents of function/risk theory suggest that consumers view adopting detection behaviors as more risky than not adopting them. However, as Levin, Schneider, and Gaeth (1998) note, this assumes that consumers ignore the long-term risks of not being screened (e.g., allowing a tumor to grow) and focus solely on short-term risks. The finding that consumer disease anxiety tends to increase screening behavior, not decrease it (McCaul et al. 1996), suggests that many consumers adopt a longer-term perspective on detection behaviors.

Second, although some detection products have short-term risks, so do many prevention products. Many people avoid flu shots for fear of pain and side effects (Centers for Disease Control and Prevention 1995); many women do not push condom use for fear of partners' reactions (Darroch and Frost 1999). If consumers focused only on short-term outcomes, they would view these preventive products as risky.

Finally, there is little empirical evidence that risk perceptions vary with product function. Several studies have examined consumer response to products differing in function (e.g., Rothman et al. 1999; Rothman et al. 1993), but none has shown that detection and prevention products differ in perceived risk.

Effects of Message Framing on Consumer Response to Product Risk

Another tenet of function/risk theory is that loss-framed messages make consumers more likely to adopt high-risk (e.g., detection) products, whereas gain-framed messages increase the attractiveness of low-risk (e.g., prevention) products. In reviewing previous framing studies, Rothman and Salovey (1997) note that loss-framed messages seem to be more persuasive in studies involving detection products, whereas gain-framed messages seem to be more effective in studies involving prevention products. To explain this apparent difference, they note that prospect theory (Tversky and Kahneman 1981) suggests that consumers prefer riskier options when outcomes are framed as losses but prefer low-risk options when outcomes are framed as gains (see also Meyerowitz and Chaiken 1987, p. 501). Thus, they conclude that loss-framed messages will persuade people to adopt detection (and other risky) products and that gain-framed messages will persuade people to adopt prevention (and other low-risk) products.

This hypothesis is compelling and, if valid, has important marketing implications. It suggests that marketers promoting prevention products (e.g., vaccinations) should use gain-framed messages, whereas those promoting detection behaviors (e.g., HIV testing) should use loss-framed messages. More broadly, it suggests that any product that consumers perceive as risky is most effectively promoted with loss-framed messages.

This theory has become increasingly influential among both persuasion researchers and practitioners. Rothman and Salovey's (1997) article has been cited more than 100 times (Thomson 2004), and the function/risk hypothesis is becoming established doctrine among social marketers (see, e.g., Arkes 2004; Ferguson et al. 2003). However, despite increasing acceptance of this hypothesis, there are unresolved questions about both its empirical support and its theoretical foundation. First, although some studies find loss framing more persuasive in promoting detection and find gain framing more persuasive in promoting prevention, other studies find just the opposite. For example, Robberson and Rogers (1988), Homer and Yoon (1992), and Blanton and colleagues (2001) all find loss framing more effective in promoting prevention behaviors (exercise, mouthwash, and condoms, respectively).

Second, in most studies cited to support function/risk theory, function is confounded with other product traits. For example, mammograms and sunscreen differ not only in function but also in the problems they address, cost, discomfort, and so forth. If framing effects differ between these two products, which of these product traits accounts for the difference? The only study to manipulate product function while holding other product attributes constant (Rothman et al. 1999, Study 1) does not find the predicted two-way interaction between function and framing.

Third, the function/risk hypothesis has theoretical problems. Assume that consumers view detection as high risk and prevention as low risk. Does the function/risk hypothesis then follow logically from prospect theory? Indeed, Tversky and Kahneman (1981) find that when outcomes were framed as losses, consumers preferred the riskier of two options. However, Tversky and Kahneman use "risk" to mean increased variance of potential outcomes; therefore, the riskier options in their experiments had both greater upside potential and greater downside potential than the no-risk (or certain) options. Given this definition of risk, choosing a risky option seems more reasonable. Prospect theory posits that people abhor a certain loss (e.g., with Option A, there is a definite loss of $2,000) and thus are so attracted to the upside potential of complete loss avoidance (e.g., Option B provides a 50% chance to lose nothing) that they are willing to tolerate the downside potential of even greater losses (e.g., a 50% chance of losing $4,000).

In contrast, when function/risk theorists argue that detection behaviors are risky, they use risk to mean an increased possibility of negative outcomes (or what Tversky and Kahneman call downside risk). Consumers who adopt detection behaviors are said to expose themselves to the possibility of short-term psychological loss (e.g., learning they have cancer), whereas nonadopters completely avoid such short-term loss potential. Thus, if the function/risk theorists' assumption that consumers focus solely on short-term outcomes is accepted, adoption of a detection behavior must be viewed as a kind of psychological Russian roulette, in which consumers actually seek potential short-term losses, with little thought to detection's long-term benefits. Such loss-seeking behavior is antithetical to the tenets of prospect theory.

Alternative Theories of Framing and Risk: Valence Congruency and Risk Discrimination

Whereas function/risk theory posits that consumers exposed to loss-framed messages actually seek short-term product risks, we propose two alternative theoretical formulations that suggest essentially the opposite: Loss-framed messages induce risk aversion, whereas gain-framed messages increase consumers' tolerance of at least some types of product risk. These two theories (valence congruency and risk discrimination) both rest on two key findings.

First, there is evidence that message framing influences consumers' moods, particularly positive moods. In depth interviews, Cox and Cox's (2001) respondents reported positive, optimistic feelings after exposure to gain-framed messages that promoted mammography and less positive moods after exposure to equivalent loss-framed messages. Similarly, Meyerowitz and Chaiken (1987, p. 506) find that gain-framed respondents were significantly more likely to report feeling "reassured." Thorsteinson and Highhouse (2003) find that respondents exposed to gain-framed want ads reported more positive scores on the Associative Valence index (e.g., "hopeful," "confident," "optimistic") than those exposed to loss-framed messages. Perhaps this finding is not too surprising, because gain-framed messages depict happy endings (in which people do the right thing and, consequently, win money, live long lives, and so forth) and loss-framed messages depict unhappy endings (in which people miss opportunities, lose money, get sick, and so forth.) However, the potential emotional impact of message framing has received remarkably little attention in the framing literature, which has tended to adopt the cognitive emphasis of behavioral decision theory (cf. Loewenstein et al. 2001).

Second, considerable research suggests that consumers exposed to affectively positive stimuli are more willing to tolerate at least some product risk to take advantage of product benefits. However, this literature contains two schools of thought that disagree on whether positive moods increase consumer tolerance of all kinds of risk or simply make consumers more likely to disregard relatively small "low-stakes" product risks.

The valence congruency theory suggests that consumers in positive moods tend to give less decisional weight to all negative product information (e.g., product risks) than consumers in negative moods. This hypothesis is often explained in terms of information priming, in which prior exposure to affectively negative material is theorized to increase the accessibility of potential negative consequences (i.e., risk) in memory, whereas exposure to affectively positive material decreases accessibility of potential negative consequences. For example, Pietromonaco and Rook (1987, p. 400) theorize that negative affect "may influence the perceived likelihood of unpleasant events because it increases the accessibility of previous similar experiences and thus facilitates imagining other potential negative outcomes."

Several studies have produced results that appear to be consistent with the valence congruency hypothesis, such that respondents in positive moods exhibit greater risk tolerance (e.g., Deldin and Levin 1986; Pietromonaco and Rook 1987; Yuen and Lee 2003). However, as Isen (2000) observes, many of these studies presented behavioral risks that were relatively minor (e.g., temporary or minor embarrassment, a small or hypothetical bet). In contrast, in tasks involving high stakes, respondents in positive moods often exhibit considerable caution regarding potential risks (see, e.g., Isen and Geva 1987). Therefore, Isen (2000) and Isen and Geva (1987) propose a more complex theory of affect and risk taking, which can be labeled risk discrimination theory. According to this theory, consumers in positive moods are better able to discriminate between important, high-stakes risks and more minor, low-stakes risks. Therefore, consumers in positive moods are more likely to disregard relatively trivial risks than are consumers in negative moods, but they may be reluctant to take large risks. In a sense, consumers in good moods are hypothesized to weigh risks more appropriately in proportion to their relative importance.

Several studies appear to support this theory (for a review, see Isen 2000). For example, in an experiment involving gambling behavior, Isen and Geva (1987, p. 145) find that "when stakes were high,... positive affect subjects were more cautious, whereas when stakes were low, positive affect subjects were more prone to take risks." Several writers have suggested that this phenomenon is part of a broader pattern, in which respondents in positive moods are simply better decision makers (i.e., they are better able to weigh and integrate information rationally and to distinguish the relevant from the irrelevant). For example, Isen (2000, p. 423) states that "positive affect promotes effective thinking about even negative material, if doing so is useful and necessary, even though it also leads people to sidestep unnecessary consideration of undesirable material. This is yet another way in which positive affect promotes flexible problem solving."

In contrast, negative affect is hypothesized to inhibit appropriate weighing of costs and benefits. In the words of Leith and Baumeister (1996, p. 1251), "negative affect apparently inhibits people from considering the affective consequences of possible actions ... [and may induce] poor choices in situations in which there is a long-term [versus] short-term tradeoff."

Hypotheses

Our first two hypotheses challenge two central tenets of function/risk theory: (1) that consumers perceive detection products as inherently riskier than prevention products and (2) that product function (detection versus prevention) moderates the persuasive effects of message framing. As we discussed previously, the theoretical and empirical bases for these tenets appear questionable. Therefore, our first two hypotheses are null or "theory refutation" hypotheses:

> H_1: Product function per se does not influence perceived product risk; that is, consumers do not perceive disease detection products as inherently riskier than prevention products that have the same product form, usage experience, and side effects.
>
> H_2: Product function per se does not moderate the persuasive effects of message framing; that is, message-framing effects do not differ between detection products and prevention products that have the same product form, usage experience, and side effects.

The next two hypotheses are based on our discussion of the valence congruency and risk discrimination theories. As we noted previously, both theories predict that message framing moderates how consumers respond to potential product risks but not in the manner predicted by function/risk theory. Consistent with both the valence congruency and the risk discrimination theories, but contrary to function/risk theory, we hypothesize the following:

> H_3: Consumers exposed to gain-framed messages exhibit less aversion to temporary product risks (i.e., those involving the potential for temporary negative effects) than consumers exposed to loss-framed messages. Consumers exposed to gain-framed messages respond less negatively to a disclosure of such temporary product risk and give less weight to perceived product risk in forming overall product evaluations than consumers exposed to loss-framed messages.

However, as we noted previously, the valence congruency and risk discrimination theories diverge on how framing influences the relative weight that consumers give to minor product risks (e.g., the risk that product use may cause temporary unpleasant effects) versus more serious product risks (e.g., the risk that product use may cause permanent harm). Valence congruency theory suggests that exposure to positive, gain-framed information should decrease the weight given to all product risks, large and small. In contrast, risk discrimination theory suggests that exposure to gain-framed messages should enhance consumers' ability to assess the emotional impact of different types of risks accurately and therefore cause them to give greater weight to serious risks than to smaller risks. Because risk discrimination has been better able to explain the range of previous empirical results, we base H_4 on this theory:

> H_4: Message framing influences the relative weight that consumers give to product risks involving temporary versus permanent adverse consequences. Consumers exposed to gain-framed messages give significantly more decisional weight to permanent product risks than to temporary product risks. In contrast, consumers exposed to loss-framed messages respond equally negatively to both types of product risk.

Study 1

Product Context and Participants

We examined our first three hypotheses in the context of skin cancer detection and prevention products. Skin cancer is the most common and one of the most rapidly growing forms of cancer. In 2002, there were 54,200 cases of skin cancer in the United States alone, and 7600 Americans died from the disease (American Cancer Society 2004).

A primary target audience for promoting skin cancer prevention and detection behaviors is young adults, a group that is especially likely to engage in behaviors (e.g., sun bathing, other outdoor recreational activities) that increase lifetime skin cancer risk. Therefore, the participants for this study were young adults whom we recruited from undergraduate business classes at a large urban university in the midwestern United States.

Experimental Design and Stimuli

In designing this experiment, we attempted to minimize some of the methodological problems that have plagued prior research on product function and risk. Specifically, in most prior studies, product function has been confounded with other product differences, such as cost, discomfort, frequency of use, and so forth. To address this problem, we developed gain- and loss-framed print advertisements for two product concepts that differ in function (detection versus prevention) but address the same underlying problem (skin cancer) and involve equivalent usage experience.

Pretest

Purpose. In the pretest, we attempted to test participants' comprehension of the message-framing and product function message components and to confirm that message framing serves as a mood manipulation. Perdue and Summers (1986) suggest that such manipulation checks are best conducted in a separate pretest rather than included in the "main" experiment that examines the dependent variables (e.g., brand attitudes, intentions). They note (p. 319) that "including these checks in the main experiment can present problems independent of whether they come before or after the dependent variable measures." Manipulation checks placed before the dependent measures may bias the latter (e.g., by creating demand effects), whereas manipulation checks placed after the dependent measures may fail to detect stimulus effects that are inherently transient, such as mood inductions (see, e.g., Vastfjall 2002). Therefore, we conducted a separate pretest to test the effects of our manipulations.

Method. We recruited 70 participants (37 male and 33 female) from undergraduate business classes. Each participant was assigned randomly to one of four experimental advertisements, which varied in behavioral function (prevention versus detection) and message framing (gain framed versus loss framed). Each advertisement had the same picture (a stylized picture of the sun) and headline ("Important News for Young Adults About Skin Cancer"). The body of each advertisement contained three components: (1) basic information about skin cancer (drawn from the American Cancer Society Web site), which was identical for all experimental groups; (2) one of two skin cancer protection product concepts that differed in function (prevention versus detection) but had the same product form and usage experience (both were lotions applied in a doctor's office that were described as "highly effective and very safe"); and (3) an advocacy message that urged consumers to adopt the product, which was either gain framed or loss framed.

After message exposure, pretest participants responded to manipulation checks for perceived product function, message framing, and mood. As a framing manipulation check, respondents used seven-point Likert scales to report whether they believed the message "emphasized the benefits of using the product" or "emphasized the costs of not using the product." As a product function manipulation check, participants reported agreement that "the lotion helps prevent skin cancer" and "the lotion helps detect skin cancer."

We measured postexposure mood using scales we adapted from the work of Howard, Gengler, and Jain (1995) and Meyerowitz and Chaiken (1987). On seven-point Likert scales, each respondent reported the degree to which 14 adjectives (e.g., assured, cheerful, afraid, happy, troubled) described his or her mood state. These responses were entered into a principal components analysis with a Varimax rotation. After the elimination of some items because of high cross-loadings, the final analysis extracted two interpretable factors with eigenvalues greater than one, which together explained 75% of the variance in the individual items. The first factor captured participants' hopefulness or optimism; items loading on this factor were "hopeful," "cheerful," "assured," and "happy." The second factor tapped into participants' anxiety mood state; items loading on this factor were "afraid," "disturbed," "troubled," and "worried." All loadings were greater than .75, and all cross-loadings were less than .15. Next, we constructed two summed scales based on the items loading on these two factors. The hopefulness mood scale had a coefficient alpha value of .84, and the anxiety mood scale had an alpha value of .92. These two summed scales were not significantly correlated with each other, suggesting that hopefulness and anxiety represent two distinct emotional dimensions, not simply different ends of the same continuum.

Pretest results. When asked whether the message communicated the "benefits of product use," participants exposed to the gain-framed version were slightly more likely to agree (X = 5.58) than those exposed to the loss-framed version (X = 5.07); however, this difference did not reach statistical significance (F = 2.69, $p = .11$). However, when asked more specifically if the advertisement "emphasized the costs of not using the product," participants in the loss-framed condition were much more likely to agree (X = 5.27) than those in the gain-framed condition (X = 4.06; F = 11.9, $p < .001$). This latter finding indicates that the participants comprehended the framing of the messages.

Participants exposed to the prevention product concept were much more likely to agree that the "lotion helps prevent skin cancer" (X = 4.90) than those exposed to the detection product concept (X = 3.11; F = 15.5, $p < .001$). Likewise, participants exposed to the detection product were much more likely to agree that the "lotion helps detect skin cancer" (X = 5.97) than those exposed to the prevention product (X = 3.04; F = 60.0, $p < .001$).

The pretest also confirmed that consumers experience more positive moods after exposure to gain-framed messages than after exposure to loss-framed messages. Participants exposed to the gain-framed message reported significantly higher levels of hopefulness (X = 3.88) than those exposed to the loss-framed message (X = 2.98; F = 8.9, $p = .004$). There was no significant difference in the reported anxiety level of the two message-framing groups.

Main Experiment

Purpose of study. This experiment has three main purposes. First, it tests H_1, that contrary to function/risk theory, disease detection products are not perceived as inherently riskier than prevention products involving the same disease, product form, and usage experience. Second, it tests H_2, that the persuasive effects of message framing do not vary with product function. As we noted previously, function/risk theory posits that loss-framed messages are most persuasive when advocating detection behaviors, whereas gain-framed messages are most persuasive when advocating prevention behaviors. However, we argue that the theoretical rationale for this hypothesis (ostensibly based on prospect theory) is problematic and that the empirical evidence in support of it is inconclusive (in part because in most studies, product function is confounded with other product differences). Thus, this experiment attempts to provide a rigorous test of the proposed interaction between function and framing, independently manipulating message frame and product function while holding other variables constant. Third, the experiment examines H_3, that consumers exposed to gain-framed messages exhibit less aversion to temporary product risks (i.e., those involving the potential for temporary negative effects) than consumers exposed to loss-framed messages. To test this hypothesis, the experiment manipulates product risk independent of product function.

Sample. We recruited 282 young adults from undergraduate business classes at a large urban university in the Midwest. The sample was 41% female and had a median age of 22.5 years.

Design and stimuli. The experiment used a 2 × 2 × 2 between-subjects factorial experimental design. The factors were message framing (gain framed versus loss framed), product function (prevention versus detection product), and product risk disclosure (presence versus absence of temporary side-effect risk). The manipulation of framing and function were identical to the pretest. We varied product risk disclosure by having the messages state either that the lotion is "very safe" or that the lotion was "generally safe" but in some patients could cause a temporary skin rash that could last for several weeks. Because a product's side-effect profile (i.e., its potential for unintended physical discomfort or harm to the product user) is conceptually distinct from its function (its core purpose or stated customer benefit), this manipulation enabled us to vary product risk independent of product function.

We randomly assigned participants to one of eight experimental advertisements. Each advertisement had the same picture, headline, and basic information about skin cancer as described in the pretest. These were followed by two paragraphs, which contained the experimental manipulations: (1) a paragraph that described one of four skin cancer protection product concepts (a low-risk prevention product, a risky prevention product, a low-risk detection product, or a risky detection product) and (2) a message, which was either gain framed or loss framed, that urged consumers to adopt the product (examples of the experimental messages appear in the Appendix). Test booklets containing the eight experimental messages were randomly assigned to participants within each classroom.

Measures. We assessed perceived product risk by having participants report their agreement on seven-point Likert scales with a series of statements regarding the potential negative consequences of product use: "Using this lotion is risky," "This lotion can lead to bad results," "Using this lotion would make me feel anxious," and "Using this lotion would cause me to worry." We combined participants' scores on these four items to form a mean scale with a coefficient alpha of .77. We assessed overall product evaluation by having participants report their agreement on seven-point Likert scales with the following statements: "Overall, I think this lotion is a very good product," "People who use this lotion will be safer than those who don't," "I would rather use this lotion than any currently available alternative," and "Using this lotion would be important to me." We combined these items to form a mean scale with a coefficient alpha of .76. We assessed behavioral intentions by having participants report on seven-point bipolar scales, anchored by "very likely" and "very unlikely," their intention to engage in three behaviors involving the target product: asking their doctor about the product, purchasing the product, and using the product. We combined these to form a mean scale with a coefficient alpha of .90. Finally, participants provided demographic information (e.g., gender, age, education) and reported on their previous skin cancer prevention behaviors (e.g., use of sunscreen) and risk factors (e.g., sun exposure, severe sunburns, family history).

Results. We examined perceived product risk with an analysis of variance (ANOVA), in which perceived product risk was the dependent variable and framing, product risk disclosure, and function were dichotomous factors. As we predicted in H_1, there was no evidence that participants perceived detection products as inherently riskier than prevention products. Product function had no impact on perceived product risk (F = .24, p = .63). The mean perceived risk of the detection product (X = 3.60) was virtually identical to that of the prevention product (X = 3.53). Thus, we found support for H_1.

However, there was a strong main effect of the side-effect manipulation on perceived product risk (F = 24.9, p < .001). Participants in the side-effects disclosure condition perceived significantly greater product risk (X = 3.9) than those in the no-side-effects condition (X = 3.2).

Message framing did not have significant main or interactive effects on participants' perceptions of product risk level. However, as we discuss subsequently, message framing altered the weight that participants gave to such risk perceptions in forming overall product evaluations.

Overall product evaluations. Next, we conducted an ANOVA that evaluated the effect of message framing, product function, and product risk disclosure on participants' overall product evaluations. Consistent with H_2, there was no interaction between message framing and product function (F = .03, p = .87).

However, framing moderated the impact of the product risk disclosure on participants' overall product evaluations (see Figure 1). As we predicted in H_3, there was a statistically significant interaction between message frame and product risk disclosure (F = 11.19, p = .001). To understand this significant interaction better, we analyzed the simple effects of the product risk disclosure in each of the two framing conditions (see, e.g., Keppel 1982, pp. 214–19). This analysis revealed that consumers exposed to loss-framed messages rated the risky (temporary side effects) product significantly more negatively (X = 4.08) than the low-risk (no side effects) product (X = 4.77; F = 13.17, p < .001). In contrast, in the gain-framed condition, consumers' product evaluations did not differ between the risky product (X = 4.40) and the low-risk product (X = 4.20; F = 1.20, p = .27).

To examine further how message framing affects consumers' sensitivity to product risk, we examined whether the correlation between perceived product risk and overall product evaluations varied significantly between the gain- and loss-framed conditions (for a similar analysis, see Pietromonaco and Rook 1987). This analysis revealed that among participants exposed to loss-framed messages, overall product evaluations were negatively correlated with perceived product risk ($r = -.31$, $p < .001$). However, among participants exposed to gain-framed messages, overall product evaluations were not significantly correlated with perceived product risk ($r = -.07$, $p = .38$). A Fisher r-to-z transformation (Kleinbaum and Kupper 1978, pp. 106–107) confirmed that the correlation between perceived product risk and overall product evaluation was significantly different across the two framing conditions ($z = 2.04$, $p < .05$). This analysis provides further evidence that participants exposed to gain-framed messages gave less evaluative weight to the product's risks than did those exposed to loss-framed messages.

Behavioral intention. Next, we conducted an ANOVA in which behavioral intention was the dependent variable and message framing, product function, and product risk disclosure were dichotomous factors. As we predicted in H_2, there was no interaction between product function and message framing ($F = 1.92$, $p = .17$). However, there was a significant interaction between product risk disclosure and message framing ($F = 4.80$, $p < .05$). Tests of simple effects revealed that participants in the loss-framed condition reported significantly greater intention to adopt the low-risk product ($X = 4.50$) than the risky product ($X = 3.75$; $F = 6.58$, $p = .01$). Among participants in the gain-framed condition, there was no significant difference in behavioral intention toward the risky ($X = 4.17$) and low-risk ($X = 4.03$) products ($F = .25$, $p = .62$) (see Figure 2).

Next, we conducted mediation analysis to determine whether within the loss-framed condition, the effect of risk disclosure on behavioral intention was mediated by overall product evaluation. We used Baron and Kenny's (1986) mediation test method, estimating a series of regression models and testing the significance of the (unstandardized) regression coefficients. First, a bivariate regression model confirmed that risk disclosure had a significant impact on the overall product evaluation ($B = -.693$; $t = -3.68$, $p < .001$). Second, a bivariate regression model confirmed that risk disclosure had a significant effect on behavioral intention ($B = -.745$; $t = -2.5$, $p = .01$). Third, we regressed

FIGURE 1
Study 1: Effect of Stated Product Risk on Overall Product Evaluations: Gain-Framed Versus Loss-Framed Conditions

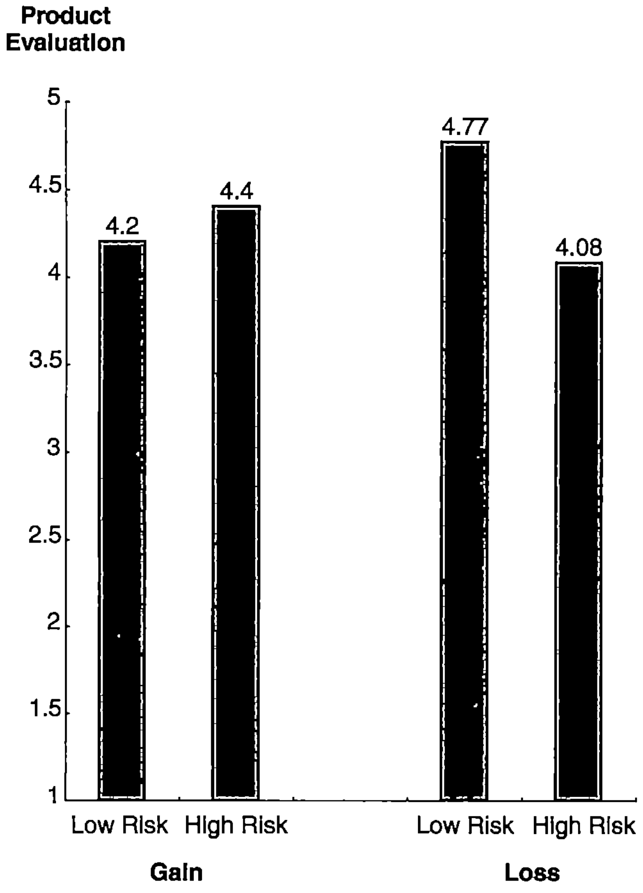
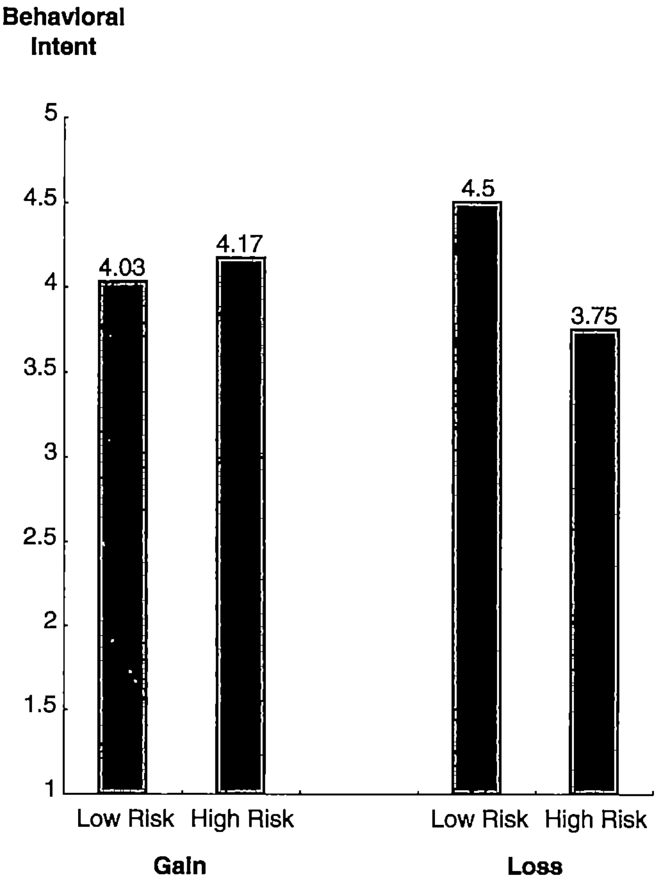


FIGURE 2
Study 1: Effect of Stated Product Risk on Behavioral Intention: Gain-Framed Versus Loss-Framed Conditions



behavioral intention simultaneously on both product risk disclosure and product evaluation. This analysis revealed that the proposed mediator (product evaluation) had a significant influence on behavioral intention (B = .986; t = 9.35, $p < .001$), whereas the effect of the independent variable (risk disclosure) became nonsignificant (B = –.06; t = –.25, $p = .80$) when the mediator was controlled. Thus, our analysis meets all four of Baron and Kenny's criteria to establish that the effect of risk disclosure on behavioral intention is completely mediated by product evaluation. Finally, Sobel's (1982) test confirmed the statistical significance of the indirect effect of risk disclosure on intention mediated by product evaluation (z = –3.43, $p < .001$).

Discussion of Study 1

Study 1 shows that consumers exposed to loss-framed messages exhibit a greater aversion to a temporary product risk (i.e., involving potential temporary negative outcomes) than do consumers exposed to gain-framed messages (who gave essentially zero decisional weight to such product risks). This finding is contrary to function/risk theory, which states that loss-framed messages induce risk-seeking behavior, but is consistent with both valence congruency theory and risk discrimination theory. However, Study 1 does not indicate whether gain-framed messages increase consumers' tolerance for all risks (as would be predicted by valence congruency theory) or simply minor, temporary risks (as predicted by risk discrimination theory). Not only is this difference theoretically interesting, but it may also have important marketing and public policy implications. If gain-framed messages cause consumers to disregard all product risk information, such messages might (intentionally or unintentionally) result in imprudent decisions about products entailing serious risks (which, as we noted previously, are increasingly being marketed directly to consumers). Conversely, if gain-framed messages merely cause consumers to disregard minor, temporary risks in the pursuit of long-term health benefits, such messages might actually help consumers make more rational decisions to protect their long-term health.

Given the importance of this issue, we address it in a second study, which examines the impact of message framing on consumers' responses to two kinds of product risk: the risk that product use may cause temporary unpleasant effects versus the risk that product use may cause permanent serious harm.

Study 2

Purpose

Study 2 serves two main purposes. First, it enables us to replicate the test of H_3 (that consumers exposed to gain-framed messages give less weight to temporary product risks than consumers exposed to loss-framed messages) in a completely different product context. Second, it enables us to test H_4 (that message framing influences the relative weight that consumers give to product risks involving temporary versus permanent adverse consequences).

Product Context

This study examines the effects of message framing and perceived product risk in the context of hepatitis B virus (HBV) vaccination adoption. Hepatitis B is a serious viral infection of the liver, which can lead to chronic hepatitis and, in some cases, death. It is transmitted in a manner similar to HIV (e.g., through unprotected sex, sharing of needles). However, the risk of HBV infection (unlike HIV) can be greatly reduced by the use of an effective vaccine, ideally administered in a series of three doses (Krugman and Stevens 1994). Unfortunately, many sexually active adults decline the vaccine, even when it is offered for free (Weinstock et al. 1995). Therefore, increased adoption of the HBV vaccine is a major social marketing goal in both the United States and other countries.

Overview of Study Design

The study is a randomized experiment. We recruited volunteers from the patient populations at three public sexually transmitted disease (STD) clinics located in two large midwestern cities, and we randomly assigned them to treatment conditions that included either gain-framed or loss-framed messages advocating the HBV vaccination. Because HBV is an STD, clinic patrons are a prime target audience for promoting the HBV vaccination.

The randomization was executed and the experimental messages and questionnaire were administered by means of an audio computer-assisted self-interview (A-CASI) tool, which is an effective tool particularly for addressing sensitive topics such as STDs (see, e.g., Turner et al. 1998.)

Before participants were exposed to the experimental messages, they were asked a series of questions about their perceptions of two distinct components of product risk: the temporary product risks associated with getting shots (e.g., pain) and the more permanent risks associated with the HBV vaccine (e.g., the possibility that the vaccine might harm their health). Then, following exposure to either gain-framed or loss-framed messages, participants were asked about their behavioral intentions regarding the product.

Participants and Procedure

A total of 213 adult (age 18 years and older) participants were recruited by a trained project manager from the patient populations of the three STD clinics over a period of several months. Participants were screened to ensure that they had no prior self-reported history of HBV infection or vaccination and that they were fluent in English. The sample was 60% male and had a median age of 27.5 years. Approximately 85% of the sample had at least a high school diploma or equivalent.

After screening for study eligibility, each participant signed a written informed consent form and then was directed to a private area with a touch-screen notebook computer with headphones. The project manager familiarized the participant in the use of the A-CASI equipment and allowed the participant to complete the interviewing process in private. The project manager had no awareness of either the message treatment group to which participants were randomly assigned or participants' questionnaire responses.

Before the framing intervention, participants completed a series of questions about demographics, prior behaviors that might affect susceptibility to HBV infection (e.g., sexual behaviors), and perceptions of both the temporary risks of HBV vaccination (e.g., injection pain) and the risks of more permanent negative outcomes (e.g., the possibility that the vaccine itself might cause HBV infection; we discuss these measures in greater depth subsequently). Next, each participant was randomly assigned a message that advocated the adoption of the HBV vaccination that was either gain framed or loss framed. After exposure to the message, participants reported their behavioral intentions.

Stimuli

The message opened with a gain- (loss-) framed general statement about the consequences of getting (or not getting) the vaccination, followed by two gain- (loss-) framed anecdotes that illustrated these consequences. To increase the perceived personal relevance of the messages, the A-CASI program automatically matched the gender (i.e., names) of the people described in the two anecdotes to the participant's gender; otherwise, the male and female messages were identical. The male version of the message is as follows (the gain-framed version appears in normal text, and the loss-framed language appears in bold):

> People who get (**don't get**) the hepatitis B shot are gaining (**losing**) a chance to protect themselves and the ones they love.
>
> Michael felt very healthy and didn't believe he would be infected by the Hepatitis B virus, [b]ut he followed (**so he didn't follow**) the advice of the American Medical Association and got (**didn't get**) 3 hepatitis B vaccine shots over 6 months. Because of this, he is (**not**) protected against hepatitis B infection. Now he is much less (**much more**) likely to get sick or die from serious liver disease, and he can look forward to (**may miss out on**) a long and healthy life, watching his child grow up.
>
> James decided he needed (**didn't need**) to protect himself and got (**didn't get**) the hepatitis B vaccine. As a result, he did not get infected (**got infected**) after having sex with an infected partner. Because James did not get infected (**got infected**) with the hepatitis B virus, he did not pass it on (**passed it on**) to his new partner. So his new partner did not get sick (**got sick**) with hepatitis B, and now they have gained (**lost**) a chance for a healthy future together.

We checked this message-framing manipulation in a pretest, in which 38 young adults (average age of 29 years) were randomly assigned either the gain- or the loss-framed message and then responded on Likert scales (7 = "strongly agree," 1 = "strongly disagree") to two statements: "This message focuses on the benefits of getting the hepatitis B shot," and "This message focuses on the negative consequences of not getting the hepatitis B shot." As we expected, participants assigned to the gain-framed message condition (X = 5.23) were more likely than the participants in the loss-framed message condition (X = 3.25) to agree that the message focused on the benefits of compliance (t = 4.29, $p < .001$). Conversely, participants in the loss-framed message condition (X = 6.25) were more likely than participants in the gain-framed message condition (X = 4.23) to agree that the message stressed the negative consequences of noncompliance (t = 3.57, $p = .001$).

Constructs and Measures

Perceived product risk. Consumers' decisions to get vaccinated are sometimes influenced by their perceptions of two fairly distinct types of product risk. First, there is the perceived risk of temporary pain or discomfort from the needle or injection itself, which sometimes deters consumers from getting vaccinations and other medical procedures involving the use of needles (see, e.g., French and Painter 1994; Zimet et al. 2004). Second, there is the perceived risk of more permanent harm caused by the specific vaccine (e.g., the perceived risk that the vaccine will transmit the disease in question or otherwise cause serious illness or long-term disability; see, e.g., Ball, Evans, and Bostrom 1998).

Before exposure to the framed messages, participants completed a series of measures that assessed these two dimensions of product risk. To assess the perceived risk of shots/needles, participants reported their agreement (on five-point Likert scales) with four items (e.g., "Shots are very painful," "Needles don't bother me at all [reversed]"), which we combined to form a mean scale with a coefficient alpha of .80. To assess perceived risk of the HBV vaccine, participants reported their agreement with six items (e.g., "The vaccine shots for hepatitis can make people very sick," "You can get infected with hepatitis from the hepatitis vaccine shots"), which we combined to form a mean scale with an alpha of .78. As we previously discussed, these two dimensions of perceived product risk appear to be distinct. The two mean scales were not significantly correlated (r = .098). Furthermore, when all ten items are combined in a factor analysis with an orthogonal rotation, the respective sets of scale items load together on two distinct factors (eigenvalues > 1) with factor loadings between .6 and .9 and minimal cross-loadings (all < .08).

Behavioral intentions. We assessed participants' intentions regarding the HBV vaccination with three items: a dichotomous yes/no question ("If you were offered the hepatitis B vaccine for free at this clinic, would you get the shot?") and two five-point Likert scale items ("I would recommend that my partner get hepatitis vaccine shots," and "I would recommend that my friends or family members get hepatitis vaccine shots"). These three items were highly intercorrelated and loaded together on a single factor that explained 78% of their combined variance. Therefore, the three items were all converted to standardized scores (to adjust for scaling differences) and then combined to form a mean scale with a coefficient alpha of .87.

Results

We tested H_3 to examine whether (as in Study 1) message framing influences the decisional weight that consumers give to temporary product risks. To address this issue, we calculated the correlation between behavioral intention and perceptions of temporary product risk (i.e., risk of short-term injection discomfort), and then we examined whether this correlation differed between the gain- and the loss-framed conditions (see Pietromonaco and Rook 1987). As in Study

1, this analysis revealed that among participants exposed to loss-framed messages, overall product evaluations were negatively correlated with perceived temporary product risk ($r = -.30$, $p < .01$). However, among participants exposed to gain-framed messages, overall product evaluations were not significantly correlated with perceived temporary product risk ($r = -.01$, $p = .92$). A Fisher r-to-z transformation (Kleinbaum and Kupper 1978, pp. 106–107) confirmed that the correlation between perceived temporary product risk and overall product evaluation was significantly different across the two framing conditions ($z = 2.18$, $p < .05$).

As an additional test of H_3, we conducted an ANOVA to examine how framing moderates consumer response to high versus low levels of temporary and permanent product risk. This analysis involved three steps. First, as we noted previously, a factor analysis of the perceived risk items produced two orthogonal factors, one representing perceived temporary product risk (i.e., shot/needle pain) and the other representing perceived permanent product risk (i.e., serious illness caused by the vaccine). Second, we performed a median split on each of these factor scores, which resulted in two uncorrelated dichotomous factors: perceived temporary product risk (high versus low) and perceived permanent product risk (high versus low). Third, we performed a 2 × 2 × 2 ANOVA of behavioral intentions, in which message framing (gain versus loss), temporary product risk (high versus low), and permanent product risk (high versus low) were dichotomous factors. This analysis showed significant, negative main effects of both permanent product risk ($F = 25.4$, $p < .001$) and temporary product risk ($F = 4.1$, $p < .05$) on behavioral intentions. However, consistent with H_3, the ANOVA also revealed a significant interaction between framing and temporary product risk ($F = 4.08$, $p < .05$). As we hypothesized, perceived temporary product risk had a significant, negative effect on behavioral intentions among loss-framed participants ($F = 6.34$, $p = .01$) but had no impact among gain-framed participants ($F = .09$, $p = .77$).

Thus, both the HBV study and the skin cancer experiment, though using different participant populations, settings, and products, produced remarkably similar results: In both studies, participants exposed to loss-framed messages gave considerably more decisional weight to temporary discomfort risk (either skin rash or shot pain) than participants exposed to gain-framed messages. Indeed, gain-framed participants in both studies gave essentially zero decisional weight to such temporary product risks.

Next, we examined H_4, that message framing influences the relative weight that consumers give to temporary product risks versus permanent product risks. As we hypothesized, among participants in the loss-framed condition, behavioral intentions were negatively correlated with both perceived temporary product risks ($r = -.30$, $p < .01$) and perceived permanent product risks ($r = -.35$, $p < .01$). To assess the difference between these values, we calculated Meng, Rosenthal, and Rubin's (1992) z statistic. Note that in testing H_3, we used the Fisher r-to-z transformation, which is most appropriate for comparing the same correlation across two different samples. However, Meng, Rosenthal, and Rubin's test is most appropriate for comparing the magnitude of two different correlations (e.g., $r_{\text{temprisk,intent}}$ versus $r_{\text{permrisk,intent}}$) within the same sample (e.g., loss-framed participants), in which each correlation is between a common dependent variable (intention) and alternative predictor variables (temporary risk versus permanent risk). Meng, Rosenthal, and Rubin's test found no significant difference between $r_{\text{temprisk,intent}}$ and $r_{\text{permrisk,intent}}$ among loss-framed participants ($z = .38$, $p = .70$), suggesting that loss-framed participants give roughly equal weight to permanent versus temporary risks in forming behavioral intentions.

Among gain-framed participants, behavioral intentions were (as we reported previously) uncorrelated with perceived temporary product risks ($r = -.01$). However, gain-framed participants' behavioral intentions exhibited a strong negative correlation ($r = -.47$, $p < .0001$) with perceived permanent product risks. Meng, Rosenthal, and Rubin's (1992) test indicates that the difference between these two correlations was statistically significant ($z = 4.03$, $p < .0001$).

Thus, as we predicted in H_4, consumers exposed to gain-framed messages appear to give significantly more decisional weight to permanent product risks than to temporary product risks. In contrast, consumers exposed to loss-framed messages respond equally negatively to both types of product risk.

Discussion

Consumers are increasingly called on to evaluate product risk (i.e., the possibility that product use may cause negative consequences). However, there has been relatively little research on how consumers evaluate product risk information, and the research that exists has produced contradictory findings. To help explain these contradictory findings, we hypothesized that the persuasive effects of product risk information may be moderated by message framing. Specifically, we evaluated three competing theories of how framing and product risk may interact to influence consumer product evaluations and intentions.

We evaluated function/risk theory (Meyerowitz and Chaiken 1987; Rothman and Salovey 1997), which posits that (1) loss-framed messages induce risk-seeking behavior, whereas gain-framed messages induce risk-avoidance behavior; (2) consumers perceive detection behaviors as risky and prevention behaviors as low risk (or "safe"); and (3) loss-framed messages increase consumers' motivation to use detection products, whereas gain-framed messages increase consumers' motivation to use prevention products.

None of these three premises was supported by our data. First, although manipulation checks confirmed that consumers understood the respective functions (detection versus prevention) of the two product concepts presented in Study 1, participants did not perceive the detection product as inherently riskier than a prevention product with the same form and usage experience. Second, Study 1 did not find any difference in message-framing effects between detection and prevention products. Third, function/risk theory's most basic premise, that loss-framed messages induce risk-seeking behavior, was contradicted by both Studies 1 and 2. In both experiments, participants exposed to loss-framed messages exhibited a general aversion to product risk involving both relatively short-term adverse effects (Studies 1 and 2) and more permanent harm (Study 2). There was no evidence of risk-seeking behavior among loss-framed participants in either experiment.

As we noted previously, proponents of function/risk theory claim that their hypothesis of risk seeking among people exposed to loss-framed messages is grounded in Tversky and Kahneman's (1981) prospect theory. In that case, does our empirical disconfirmation of this hypothesis mean that prospect theory is wrong? Absolutely not. It simply means that in the attempt to derive function/risk theory from prospect theory, key aspects of prospect theory (particularly the meaning of the word "risk") were lost in translation. As we noted previously, Tversky and Kahneman used the word "risk" in a technical sense to mean the variance of potential outcomes, so that a riskier option entails both extremely positive and extremely negative possibilities rather than a predictable (but mediocre) certain option. When risk is operationalized in this way, participants are sometimes motivated to choose the risky option because of its positive possibilities. This is particularly true if outcomes are framed as losses because the certain option then entails a certain small loss, whereas the risky option offers the upside possibility of no loss. Given this somewhat esoteric definition of "risk," risk-seeking behavior seems plausible.

However, although this "risk as variance" definition is used in a few academic fields, including expected utility theory (the point of departure for prospect theory) and financial investment analysis, it does not match the way the word "risk" is used by function/risk theorists or, for that matter, in most academic disciplines. To most people, academic or layman, "risk" specifically connotes the possibility of something bad. For example, *Merriam-Webster* (2005) defines risk as "possibility of loss or injury: peril." Similarly, in a typical consumer behavior definition, Solomon (2004, p. 304) defines perceived risk as "the belief that the product has potentially negative consequences," and risk perception researcher Paul Slovic (2000) uses the words "risk" and "hazard" interchangeably. In none of these cases does "risk" imply the possibility of extremely positive consequences that help explain Tversky and Kahneman's (1981) results.

Therefore, to understand how message framing might influence consumers' responses to product risk (defined as potential negative consequences), we proposed two alternative theories: valence congruity theory and risk discrimination theory. Both theories are based on prior findings that message framing influences mood and that mood influences consumers' responses to risk. Both theories suggest that exposure to gain-framed messages increases consumers' tolerance for temporary product risks, a finding that was confirmed in both Study 1 and Study 2. However, the theories diverge in predicting how message framing will affect consumers' responses to more serious, long-lasting product risks. Thus, Study 2 was designed to address this issue. The findings of this experiment support risk discrimination theory: Participants exposed to gain-framed messages appear to be better able to discriminate between important and unimportant risks; they essentially ignore minor or temporary product risks but exhibit considerable caution regarding the possibility of more permanent, serious risks.

In addition to their theoretical implications, our findings may have implications for those who design and regulate promotional messages for risky products, including DTC advertisements for prescription drugs. Food and Drug Administration requirements dictate that any DTC advertisement that mentions a product's brand name (e.g., Levitra) and its benefits (improving male sexual function) must disclose the product's side-effect risks (e.g., severe headaches). The only way to avoid such disclosure is to either not name the product (e.g., "Ask your doctor about new treatments for erectile dysfunction") or not say what it does (e.g., "Ask your doctor about Levitra"). Thus, DTC advertisers must either mention product risks or produce rather cryptic advertisements (e.g., the Levitra advertisement that does not indicate what it does but shows a football going through a swinging tire).

Although there has been relatively little research on how consumers process such product risk information, there are several theories. One theory (e.g., Goetzl 2000; Rodale Publications 1999) is that product risk disclosures actually enhance advertisement credibility by making the advertiser seem more forthcoming. Another theory is that product risk disclosures undermine the positive content of an advertisement and reduce advertising persuasiveness (Goetzl 2001; Morris, Mazis, and Brinberg 1989). However, our findings do not support either of these simple main effects of product risk. Instead, our data suggest that consumers respond differently to product risk depending on both the nature of this risk (temporary versus permanent) and the framing of the persuasive message. Participants exposed to loss-framed messages exhibited an aversion to product risk, whereas those exposed to upbeat, gain-framed messages essentially disregarded temporary product risks in forming product evaluations and intentions.

These results suggest that advertisers are able to reduce the decisional weight that consumers give to certain types of product risk information by placing such information in the context of upbeat, gain-framed message executions. If confirmed by further research, this finding has potentially important implications both for marketers required to include product risk disclosures in their promotional materials and for the consumer protection agencies that institute these disclosure requirements.

To understand the policy implications of these findings more fully, further research should explore several issues. Specifically, further research should examine whether the framing-induced optimism we observed in our study is beneficial or harmful in general to consumer welfare. As we noted previously, our research found that consumers exposed to gain-framed messages gave essentially zero weight to temporary product risks (i.e., skin rash or shot pain) in evaluating a disease protection product (i.e., skin cancer lotion or hepatitis vaccination). However, is this disregard of temporary product risk a good thing or a bad thing? On the one hand, in general, consumers should consider risks before pursuing any course of action. On the other hand, perhaps consumers should give the possibility of a skin rash (or a sore shoulder) considerably less weight than the chance to protect themselves from skin cancer or liver disease. Perhaps optimism induced by gain-framed messages simply helps consumers overcome their natural hesitancy toward taking calculated risks (i.e., the omission bias). This is an important issue for further research.

Finally, further research should explore the emotional effects of message framing, examining more broadly how these effects may influence consumer judgment and decision making. As we noted previously, most message-framing research to date has been grounded (at least ostensibly) in Kahneman and Tversky's (1979; see also Tversky and Kahneman 1981) prospect theory, which, as is the case with much of behavioral decision theory, tends to emphasize cognitive over affective phenomena (for a discussion of the neglect of emotion in behavioral decision research, see, e.g., Loewenstein et al. 2001). However, the finding that message framing influences consumers' moods, along with the wealth of research showing that mood affects decision making (e.g., Isen 2000), could open up entirely new avenues in the study of framing effects. We strongly encourage researchers to explore those avenues.

Appendix
Examples of Experimental Messages

Detection/Low Risk/Loss

Skin cancer is the most common form of cancer among people in their twenties. Every 53 minutes, someone dies of skin cancer. Skin cancer can be hard to see with the naked eye or may look like a normal mole or freckle.

Fortunately, there is now a medicated lotion which can detect skin cancer. This product is not a sunblock or sunscreen. It is applied by a doctor to the patient's skin and creates a gentle reaction, which allows the doctor to easily detect the early development of skin cancer. This lotion is highly effective and very safe.

If you fail to use this lotion, you will be losing an opportunity to detect the development of skin cancer cells before they can do serious harm. By failing to use this lotion, you can significantly harm your chances of living a long, cancer-free life.

For more information, call 1-800-4CANCER, or talk to your doctor.

Prevention/High Risk/Gain

Skin cancer is the most common form of cancer among people in their twenties. Every 53 minutes, someone dies of skin cancer. Skin cancer can be hard to see with the naked eye or may look like a normal mole or freckle.

Fortunately, there is now a medicated lotion which can prevent skin cancer. This product is not a sunblock or sunscreen. It is applied by a doctor to the patient's skin and creates a gentle reaction, which allows the doctor to easily prevent the early development of skin cancer. This lotion is highly effective and generally safe. However, in some patients, it can cause a severe skin rash lasting several weeks.

If you use this lotion, you will be gaining an opportunity to prevent the development of skin cancer cells before they can do serious harm. By using this lotion, you can significantly improve your chances of living a long, cancer-free life.

For more information, call 1-800-4CANCER, or talk to your doctor.

REFERENCES

Adaval, Rashmi (2001), "Sometimes It Just Feels Right. The Differential Weighting of Affect-Consistent and Affect-Inconsistent Product Information," *Journal of Consumer Research*, 28 (June), 1–17.

American Cancer Society (2004), "Skin Cancer Facts," (accessed November 10, 2004), [available at www.cancer.org/docroot/PED/content/ped_7_1_What_You_Need_To_Know_About_Skin_Cancer.asp?sitearea=PED].

Arkes, Hal (2004), "The Psychology of Patient Decision Making: Framing," in *Interactive Textbook of Clinical Symptom Research*, (accessed December 2, 2004), [available at http://symptomresearch.nih gov/chapter_4/sec1/chas1pag1.htm].

Ball, Leslie, Geoffrey Evans, and Ann Bostrom (1998), "Risky Business: Challenges in Vaccine Risk Communication," *Pediatrics*, 101 (3), 453–58.

Baron, Reuben and David A. Kenny (1986), "The Moderator-Mediator Distinction in Social Psychological Research: Conceptual, Strategic and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (6), 1173–82.

Blanton, Hart, R. Vanden Eijnden, B. Buunk, F. Gibbons, M Gerrard, and A. Bakker (2001), "Accentuate the Negative: Social Images in the Prediction and Promotion of Condom Use," *Journal of Applied Social Psychology*, 31 (2), 274–95.

Calfee, John (2002), "Public Policy Issues in Direct-to-Consumer Advertising of Prescription Drugs," *Journal of Public Policy & Marketing*, 21 (Fall), 174–93.

Centers for Disease Control and Prevention (1995), "Improving Influenza Vaccination Rates for Medicare Beneficiaries: Montana and Wyoming," *Morbidity and Mortality Weekly Report*. Washington, DC: U.S. Department of Health and Human Services

Cohen, Joel (2002), "Introductory Comments: Direct-to-Consumer Prescription Drug Advertising: Evaluating Regulatory Policy in the United States and New Zealand," *Journal of Public Policy & Marketing*, 21 (Fall), 172–73.

Cox, Dena and Anthony D. Cox (2001), "Communicating the Consequences of Early Detection: The Role of Evidence and Framing," *Journal of Marketing*, 65 (July), 91–103

Darroch, Jacqueline and Jennifer Frost (1999), "Women's Interest in Vaginal Microbicides," *Family Planning Perspectives*, 31 (February), 16–23.

Deldin, Patricia and Irwin B. Levin (1986), "The Effect of Mood Induction in a Risky Decision-Making Task," *Bulletin of the Psychonomic Society*, 24 (1), 4–6.

Ferguson, Eamonn, Peter Bibby, Joanna Leaviss, and Andrew Weyman (2003), *Effective Design of Workplace Risk Communication*, (accessed December 2, 2004), [available at www.hse.gov.uk/research/rrpdf/rr093.pdf].

Fitch, Malcolm (1997), "What You Can Really Learn from Fund Performance," *Money*, (March), B1–B4.

Fox, Richard, Dean Krugman, James Fletcher, and Paul Fisher (1998), "Adolescents' Attention to Beer and Cigarette Print Ads and Associated Product Warnings," *Journal of Advertising*, 27 (Fall), 57–67.

French, Gina and Eileen Painter (1994), "Blowing Away Shot Pain: A Technique for Pain Management During Immunization," *Pediatrics*, 93 (3), 384–88.

Goetzl, David (2000), "TVB Study: DTC Disclosures Lend Credibility to Drug Ads," *Advertising Age*, (June 5), 6.

——— (2001), "Xenical Reworks Its Ads; TV Spots Stress Awareness, Skip Over Side Effects," *Advertising Age*, (February 5), 3.

Homer, Pamela and Sun-Gil Yoon (1992), "Message Framing and the Interrelationships Among Ad-Based Feelings, Affect and Cognition," *Journal of Advertising*, 21 (March), 19–32.

Howard, Daniel, Charles Gengler, and Ambuj Jain (1995), "What's in a Name? A Complimentary Means of Persuasion," *Journal of Consumer Research*, 22 (September), 200–211.

Isen, Alice (2000), "Positive Affect and Decision Making," in *Handbook of Emotion*, M. Lewis and J. Haviland-Jones, eds. New York: Guilford, 417–35.

——— and Nehemia Geva (1987), "The Influence of Positive Affect on Acceptable Level of Risk: The Person with a Large Canoe Has a Large Worry," *Organizational Behavior and Human Decision Processes*, 39 (April), 145–54.

Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47 (2), 263–91.

Keppel, Geoffrey (1982), *Design and Analysis*. Englewood Cliffs, NJ: Prentice Hall.

Kleinbaum, David and Lawrence Kupper (1978), *Applied Regression Analysis and Other Mutivariable Methods*. North Scituate, MA. Duxbury Press.

Krugman, S. and C. Stevens (1994), "Hepatitis B Vaccine," in *Vaccines*, S. Plotkin and E. Mortimer, eds. Philadelphia: W.B. Saunders, 419–37.

Leith, Karen and Roy Baumeister (1996), "Why Do Bad Moods Increase Self-Defeating Behavior? Emotion, Risk-Taking and Self-Regulation," *Journal of Personality and Social Psychology*, 71 (6), 1250–67.

Levin, Irwin, Sandra Schneider, and Gary Gaeth (1998), "All Frames Are Not Created Equal· A Typology and Critical Analysis of Framing Effects," *Organizational Behavior and Human Decision Processes*, 76 (November), 149–88.

Loewenstein, G., E. Weber, C. Hsee, and N. Welch (2001), "Risks as Feelings," *Psychological Bulletin*, 127 (2), 267–86.

McCaul, Kevin, Ann Branstetter, Dawn Schroeder, and Russell Glasgow (1996), "What Is the Relationship Between Breast Cancer Risk and Mammography Screening? A Meta-Analytic Review," *Health Psychology*, 15 (6), 423–29.

Meng, Xiao-Li, Robert Rosenthal, and Donald Rubin (1992), "Comparing Correlated Correlation Coefficients," *Psychological Bulletin*, 111 (3), 172–75.

Menon, Ajit, Aparna Deshpande, Matthew Perri, and George Zinkhan (2003a), "Communicating Risk Information in Print Direct-to-Consumer Prescription Drug Advertisements: Implications for Public Health," paper presented at the 131st American Public Health Association Annual Meeting and Exposition, San Francisco (November 15–19), (accessed December 2, 2004), [available at http://alpha.confex.com/apha/131am/techprogram/paper_58577.htm].

———, ———, ———, and ——— (2003b), "Consumers' Attention to the Brief Summary in Print Direct-to-Consumer Advertisements: Perceived Usefulness in Patient–Physician Discussions," *Journal of Public Policy & Marketing*, 22 (Fall), 181–91.

Merriam-Webster (2005), *Merriam-Webster Online Dictionary*, (accessed May 1, 2005), [available at http.//www.m-w.com/cgi-bin/dictionary?book=Dictionary&va=risk&x=13&y=14].

Meyerowitz, Beth and Shelley Chaiken (1987), "The Effort of Message Framing on BSE Attitudes, Intentions and Behavior," *Journal of Personality and Social Psychology*, 52 (3), 500–510.

Morris, Louis, Michael Mazis, and David Brinberg (1989), "Risk Disclosures in Televised Prescription Drug Advertising to Consumers," *Journal of Public Policy & Marketing*, 8 (Spring), 64–70.

Perdue, Barbara and John O. Summers (1986), "Checking the Success of Manipulations in Marketing Experiments," *Journal of Marketing Research*, 23 (November), 317–26.

Pietromonaco, P. and K. Rook (1987), "Decision Style in Depression: The Contribution of Perceived Risks and Benefits," *Journal of Personality and Social Psychology*, 52 (2), 399–408.

Robberson, M.R. and R.W. Rogers (1988), "Beyond Fear Appeals: Negative and Positive Persuasive Appeals to Health and Self-Esteem," *Journal of Applied Social Psychology*, 13 (March), 277–87.

Rodale Publications (1999), *Prevention Magazine: A National Survey of Consumer Reactions to Direct-to-Consumer Advertising*. Emmaus, PA· Rodale Publications.

Rothman, Alexander J., Steven C Martino, Brian T. Bedell, Jerusha B. Detweiler, and Peter Salovey (1999), "The Systematic Influence of Gain- and Loss-Framed Messages on Interest in and Use of Different Types of Health Behavior," *Personality and Social Psychology Bulletin*, 25 (November), 1355–69.

——— and Peter Salovey (1997), "Shaping Perception to Motivate Healthy Behavior: The Role of Message Framing," *Psychological Bulletin*, 121 (January), 3–19.

———, ———, Carol Antone, Kelli Keough, and Chloe Martin (1993), "The Influence of Message Framing on Intentions to Perform Health Behaviors," *Journal of Experimental Social Psychology*, 28 (September), 408–433.

Slovic, Paul (2000), *The Perception of Risk*. London· Earthscan.

Sobel, Michael E. (1982), "Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models," *Sociological Methodology*, 13, 290–312.

Solomon, Michael (2004), *Consumer Behavior*. Upper Saddle River, NJ. Pearson/Prentice Hall.

Thomaselli, Rich (2004), "Marketers Left in Limbo by DTC Delay," *Advertising Age*, (September 20), 3.

Thomson (2004), *Science Citation Index Expanded*, (accessed December 2, 2004), [available at http://www.isinet.com/products/citation/scie/].

Thorsteinson, Todd and Scott Highhouse (2003), "Effects of Goal Framing in Job Advertisements on Organizational Attractiveness," *Journal of Applied Social Psychology*, 31 (11), 2393–2412.

Turner, C., L. Ku, S. Rogers, L. Lindberg, J. Pleck, and F. Sonenstein (1998), "Adolescent Sexual Behavior, Drug Use and Violence: Increased Reporting with Computer Survey Technology," *Science*, (March 8), 867–73.

Tversky, Amos and Daniel Kahneman (1981), "The Framing of Decisions and the Psychology of Choice," *Science*, 211 (January), 453–58.

Vastfjall, Daniel (2002), "Influences of Current Mood and Noise Sensitivity on Judgments of Noise Annoyance," *Journal of Psychology*, 136 (July), 357–70.

Weinstock, H., G. Bolan, J. Moran, T. Peterman, L. Polish, and A Reingold (1995), "Routine Hepatitis B Vaccination in a Clinic for Sexually Transmitted Diseases," *American Journal of Public Health*, 85 (June), 846–49.

Yuen, K. and T. Lee (2003), "Could Mood State Affect Risk-Taking Decisions?" *Journal of Affective Disorders*, 75 (1), 11–18.

Zimet, G., S. Rosenthal, J.D. Fortenberry, R. Brady, W. Tu, J. Wu, D. Bernstein, L. Stanberry, K. Stone, K. Fife, and J. Leichliter (2004), "Factors Predicting the Acceptance of Herpes Simplex Virus Type 2 Antibody Testing," *Sexually Transmitted Diseases*, 31 (11), 665–69

Kersi D. Antia, Mark E. Bergen, Shantanu Dutta, & Robert J. Fisher

How Does Enforcement Deter Gray Market Incidence?

Gray market activity has become increasingly prevalent. The prevailing wisdom in marketing is to use more severe enforcement to deter gray marketing. However, the certainty and speed of enforcement may also have a bearing on the incidence of violations. This article examines whether and how enforcement deters gray marketing. The results from a field survey of manufacturers and an experimental design suggest that, by itself, enforcement severity has no impact. Deterrence results only when the multiple facets of enforcement are used in combination.

In today's competitive markets, manufacturers increasingly rely on their authorized distributors to perform value-adding functions (Coughlan et al. 2001). To safeguard distributor incentives to perform these services, manufacturers typically deploy resale restrictions through explicit contracts or implicit agreements. By circumscribing to whom distributors may sell, resale restrictions limit intrabrand competition and maintain distributor margins. Gray market activity—that is, the sale of genuine trademarked products through distribution channels unauthorized by the manufacturer or brand owner—poses a direct and significant threat to manufacturers' deployed resale restrictions. Given free reign, gray markets create free-riding problems across distributors that provide customer service, make a selective distribution system more intensive, and harm distributors that have made specific investments in the channel of distribution. Gray marketing is problematic for manufacturers because it can have a negative impact on distributor relations and the manufacturer's brand equity, ultimately undermining the integrity of the distribution channel (Corey, Cespedes, and Rangan 1989; Myers and Griffith 1999).

Gray markets are endemic across, though not limited to, a wide variety of industries, ranging from heavy construction equipment, personal computers, and cellular phones to perfumes, watches, personal care, and other consumer products. Estimates of gray market activity range from $10 billion in economy-wide annual gray market sales (Cespedes, Corey, and Rangan 1988) to $20 billion in the information technology sector alone (estimate by the Alliance for Gray Market and Counterfeit Abatement; see www.agmaglobal.org). Surveys conducted by Michael (1998) and Myers (1999) confirm the increasing incidence and scope of gray market activity.

Most firms are keenly aware of the costs of allowing gray market activity to go unchecked (Bucklin 1993; Hwang 1999; Nelson 1999) and allocate significant resources to limiting the incidence of violations. In general, the prevailing advice from the academic literature is that increasing enforcement severity (i.e., raising the magnitude of the punitive response undertaken) lowers the incidence of gray market activity. For example, the enforcement literature in marketing (Banerji 1990; Dutta, Bergen, and John 1994) exclusively focuses on severity as the dimension of enforcement that firms should manage. Such a focus reflects the implicit assumption that the critical management variable for manufacturers seeking effective gray market deterrence is "the severity of the principal's disciplinary response to an agent's violation of a contractual obligation" (Antia and Frazier 2001, p. 67).

Although such advice may seem intuitively appealing and finds favor with those urging manufacturers to "get down to business" against gray market activity (Nelson 1999), to the best of our knowledge, there is no evidence of the efficacy of severe enforcement in the context of gray markets. A well-established literature spanning economics, law, sociology, and social psychology, which has come to be known as the deterrence doctrine, proposes that severe enforcement alone may not be enough to curb the incidence of violations. Instead, the doctrine advocates a broader notion of enforcement, comprising (in addition to severity) the certainty and speed of the response (Howe and Brandau 1988; Manson 2001; Posner 1985).

Understanding how these dimensions of enforcement work together is essential to managing gray market incidence more effectively. Yet to the best of our knowledge, no attempt has been made to incorporate the possible role of severity, certainty, and speed of enforcement together in curbing gray market incidence; severity remains the sole

---

Kersi D. Antia is an assistant professor (e-mail: kantia@ivey.uwo.ca), and Robert J. Fisher is R.A. Barford Professor of Communications (e-mail: rfisher@ivey.uwo.ca), Richard Ivey School of Business, University of Western Ontario. Mark E. Bergen is Carlson Professor of Marketing, Department of Marketing and Logistics Management, Carlson School of Management, University of Minnesota (e-mail: mbergen@csom.umn.edu). Shantanu Dutta is Tappan Professor of Business-to-Business Marketing, Department of Marketing, Marshall School of Business, University of Southern California (e-mail: sdutta@marshall.usc.edu). This research was supported by a Social Sciences and Humanities Research Council of Canada grant (No. 410-2002-0605) at the University of Western Ontario; additional funding was provided by the Zumberge Foundation, University of Southern California, and the John M. Olin School of Business, Washington University, St. Louis. The authors gratefully acknowledge the assistance of the Beauty and Barber Supply Institute in the data collection efforts and the comments on previous drafts of this article by Rajesh Chandy, John Hulland, Jaideep Prabhu, and Sourav Ray.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

focus of any discussion of enforcement, both in the trade press (Hwang 1999; Nelson 1999) and in the relevant marketing literature (Antia and Frazier 2001; Bergen, Heide, and Dutta 1998). Accordingly, current enforcement-related advice to managers is limited.

In this article, we attempt to broaden our conceptualization of enforcement and explore the incidence of gray market violations as a function of the severity, certainty, and speed of the manufacturer's general enforcement behavior. In the interests of appropriate model specification, we also incorporate several context-specific determinants of incidence gained from multiple in-depth interviews and a review of the rich literature on gray markets (Ahmadi and Yang 2000; Cavusgil and Sikora 1988; Corey, Cespedes, and Rangan 1989; Coughlan and Soberman 1998).

We first test our hypotheses with data collected from a field survey of 104 manufacturers in the personal care products industry. The difficulty of obtaining such data is well acknowledged, given the sensitive nature of the phenomenon and its implications (Cavusgil and Sikora 1988; Myers 1999). Ours is the first study to bring evidence to bear on the relationship between enforcement and gray market incidence. We then complement the insights gained from the field study with an experiment that we designed to (1) provide the perspective of the target rather than the source of enforcement and (2) manipulate the facets of enforcement in a controlled setting. Together, both studies yield rich insights into the impact of enforcement on deterrence.

In the sections that follow, we introduce the deterrence doctrine as the lens through which we view the relationship between gray market incidence and enforcement. We then discuss hypotheses that link gray market incidence to the three facets of enforcement. We follow this with descriptions of the research method and a discussion of the results and implications of our studies. We conclude with the limitations of our study and provide avenues for further research.

Conceptual Framework

Gray Market Incidence, Enforcement, and the Deterrence Doctrine

The notion of deterrence has its roots in the seminal works of Bentham (1962) and Beccaria (1963) and refers to the preventive effect of actual or threatened punishment on potential offenders (Ball 1955). In the current context, the aim of deterrence is to minimize or preclude violations of deployed resale restrictions not only by the perpetrator but also by other potential offenders. Thus, we relate the incidence of gray market violations to the manufacturer's enforcement behavior in general rather than the punitive response to a particular violation (cf. Antia and Frazier 2001; Bergen, Heide, and Dutta 1998).

Within a gray market context, enforcement refers to the means by which the manufacturer ensures distributor compliance with resale restrictions, including fines, litigation, social ostracism, and termination. The manufacturer's ability to enforce stems from the legitimate authority bestowed by explicit or implicit agreements with its intermediaries (see French and Raven 1959), but this is not the only way to deter gray market activities or achieve dealer compliance. In addition to enforcement, manufacturers may also withhold ongoing benefits, such as promotional discounts, advertising support, and training, from noncompliant downstream intermediaries. Such influence attempts are related to, though distinct from, enforcement as we define it.

Severity is fundamental to the notion of enforcement and can be defined as the strength or intensity of corrective actions across detected violations (Gibbs 1975). The costs imposed on the violating party are a direct function of the severity of the punishment that is received. The magnitude of a fine affects the offending distributor's net payoff from a gray market violation (in this case, the profit captured by the violation less the fine). Other nonfinancial costs may also be important. The costs associated with litigation and termination include the time spent in legal consultation and/or in seeking alternative sources of supply. Gray market violations are also likely to have a negative effect on the relationship between the manufacturer and the supplier, thus increasing monitoring and surveillance costs. Overall, the potential gains from violating resale restrictions are eroded by nontrivial costs, which reduces the incentive to commit the violation (Becker 1968; Stigler 1970).

The marketing literature treats enforcement and its severity as essentially synonymous. Dutta, Bergen, and John (1994) and Bergen, Heide, and Dutta (1998) conceptualize enforcement in terms of termination. Antia and Frazier (2001) explicitly define enforcement in terms of severity. Moreover, as it currently stands, the preceding literature could be characterized as advocating (even taking for granted) a single-minded focus on severity. Such a focus rests on the assumption that the rigor of enforcement has a significant inverse relationship to the incidence of violations, notwithstanding variations in certainty or speed. Thus:

> H_1: The greater the severity of enforcement, the lesser is the likelihood of gray market incidence.

Interaction with Certainty

In contrast to the single-minded focus on severity as a determinant of deterrence in the marketing literature, work in other disciplines emphasizes the need to consider enforcement certainty (the likelihood that a violator will be punished) and speed (the time taken by the manufacturer to undertake corrective action). As originally conceptualized and specified, the deterrence doctrine emphasized the facets of severity and certainty (Andenaes 1971; Bailey and Smith 1972; Bentham 1962; Gibbs 1968). Subsequent augmentations acknowledge the potential role of speed and incorporate its impact on deterrence (Gray et al. 1982). Mirroring the development path of the literature, we first focus on the potential deterrent impact of severity and certainty and then incorporate the role of speed in our hypotheses.

Research in sociology conceptualizes deterrence as a joint function of severity and certainty (Antunes and Hunt 1973; Gray and Martin 1969). The severity of enforcement is posited to have significant deterrent value but only when it is matched with a high likelihood of action being taken.

The notion of matching facets of enforcement implies the need for high levels of each facet of enforcement in order to achieve deterrence.

By certainty, we mean the manufacturer's general propensity to undertake enforcement in response to violations (Gibbs 1975). To take corrective action, the manufacturer must be able to obtain information about violations once they occur (Dutta, Heide, and Bergen 1999) and have the motivation to enforce (Antia and Frazier 2001). This is our point of departure from the deterrence doctrine, as represented in the sociology literature. In the latter case, the motivation to eradicate or at least minimize crime is a given. In contrast, manufacturers that detect gray market violations may undertake selective enforcement for a variety of reasons (Banerji 1990; Bergen, Heide, and Dutta 1998; Coughlan and Soberman 1998). In the current context, therefore, certainty of enforcement is a function of the manufacturer's ability and motivation. The former is reflected in the manufacturer's detection ability because enforcement hinges on first knowing that a violation has occurred (Antia and Frazier 2001).

Unfortunately, gray markets raise three well-known problems with regard to inferring motivation. First, in contrast to sociology, the certainty of enforcement cannot be estimated conveniently from archival information on reported crime statistics and conviction rates (Gibbs 1975), because no such information exists for gray market violations. In the absence of archival information on the incidence of violations and manufacturers' enforcement responses, it is necessary to rely on microlevel, survey-based data to gain insights into this highly sensitive issue. Second, manufacturers do not know of all instances of violations (Wathne and Heide 2000), because detection is costly (Stigler 1970) and imperfect (Dutta, Heide, and Bergen 1999). Third, and perhaps equally important, no manufacturer would admit to selective enforcement, thus introducing social desirability bias (Fisher 1993).

Recognizing this potential for bias, survey-based studies on enforcement have not solicited informant evaluations of their own likelihood of taking action. Instead, motivation is inferred from the reported enforcement response to particular instances of violations (for recent examples of this approach, see Antia and Frazier 2001; Bergen, Heide, and Dutta 1998). Moreover, previous studies have hypothesized and found evidence of a direct and strong relationship between detection ability and likelihood of enforcement (Antia and Frazier 2001; Dutta, Bergen, and John 1994; Dutta, Heide, and Bergen 1999; Ghosh and John 1999). Given the preceding difficulty and in light of the strong relationship between the two constructs, we represent certainty in terms of detection ability. Our approach is consistent with and builds on the existing marketing literature on enforcement.

The rationale for matched and high levels of severity and certainty suggests that when perpetrators decide whether to engage in gray market activities, they consider not only the previously discussed costs imposed by severe enforcement but also the likelihood of severe enforcement. If detection ability and, consequently, the certainty of an enforcement response are low, potential miscreants are likely to discount the costs imposed by severe enforcement (Luckenbill 1982). Severity poses no threat to the offender if the offence remains undetected by the manufacturer. Similar discounting of negative outcomes occurs if, for example, high detection capabilities are accompanied by less severe enforcement behavior. In this case, potential violators may reason that though the likelihood of apprehension is high (as a result of well-honed detection capabilities), their offenses are likely to incur nothing more serious than a "slap on the wrist" (Stafford et al. 1982). Proponents of this perspective argue that the credibility of sanctions hinges on both detection ability and severity (Gray et al. 1982; Hollinger and Clark 1983). The interaction perspective is also known as the "credibility of severe sanctions hypothesis" (Grasmick and McLaughlin 1978) and suggests the following:

> H_2: Severe enforcement deters gray market incidence when the certainty of enforcement is high

The Augmented Deterrence Doctrine

Early empirical work on deterrence (Silberman 1976) focused exclusively on the certainty and severity of enforcement. Only in later studies (Gray et al. 1982) did scholars acknowledge and incorporate a third, additional dimension of enforcement: speed. The speed of enforcement refers to the time elapsed between the detection of violations and the corrective actions taken in response. Similar to severity and detection ability, our notion of speed is the time taken to respond to detected violations.

Although the relationship between enforcement speed and deterrence is not as well specified as it is for detection ability and severity, our interdisciplinary literature review yields some insights into the effects of variation in enforcement speed. First, swift (slow) corrective action reduces (increases) the length of time over which the party committing the violation may enjoy its payoff (Tedeschi 1976). Second, delay in applying sanctions allows the violating party "room to maneuver," that is, to undertake actions to avoid (or at least minimize) bearing the brunt of the costs imposed by corrective action (Hufbauer, Schott, and Elliott 1990). Third, research in psychology (Diver-Stamnes and Thomas 1995) and law (Manson 2001) suggests that temporal proximity of the violation and the consequent corrective enforcement action reinforce the punitive consequences to the perpetrator, as well as the group at large.

In contrast to the well-documented interaction between severity and detection ability, there has been little, if any, attention focused on explicating the interactions involving speed and severity (for a notable exception, see Gray et al. 1982). The logic underlying this interaction mirrors that which we proposed in H_2, that the deterrent impact of enforcement is deemed to reside in high levels of speed combined with severity. It is plausible that severe enforcement may do little to deter violations if the enforcement response is inordinately delayed. Because a slow response allows the perpetrator time to enjoy a longer payoff from the violation, gives the perpetrator room to maneuver to avoid bearing the full brunt of enforcement, and/or weakens the causal link between violation and enforcement, delayed enforcement could dampen the deterrent impact of severity. By extension, severity should have a significant effect on deterrence when both the likelihood of detection and the speed of enforcement are high. An enforcement strategy that involves high levels of all three components serves to emphasize the manufacturer's commitment to maintaining system integrity by presenting a unified, unambiguous deterrent. Thus:

H_3: Severe enforcement deters gray market incidence when the speed of enforcement is high.

H_4: Severe enforcement deters gray market incidence when both detection ability and the speed of enforcement are high.

Context-Specific Determinants of Gray Market Incidence

Scholars of deterrence agree that any examination of deterrence must account for contextual variables that have a bearing on the incidence of violations, independent of enforcement (Gibbs 1975; Meier and Johnson 1977). Our review of the academic literature on gray markets yielded five factors that emerged most frequently: price differentials, premium brand positioning, product scarcity, free-riding potential, and customer heterogeneity on services demanded.

According to Onkvisit and Shaw (1989, p. 205), "price differential is the only true reason for the gray market to exist." In attempting to price optimally for local conditions (Bucklin 1993; Duhan and Sheffet 1988), manufacturers facilitate arbitrage, whereby the diverter may source product from the low-priced market and sell it in the high-priced market without the manufacturer's authorization (Assmus and Wiese 1995). A related driver of gray market violations is the extent to which the product has a unique and differentiated brand name (Bucklin 1993). The ability to obtain a product with brand name cachet at significantly lower prices is a powerful motivator of consumer purchase (Assmus and Wiese 1995), which in turn drives the occurrence of gray marketing. Gray market violations are also likely to be high when the authorized distributors in a market are unable to satisfy demand for the product. Pent-up demand for a product creates the incentive necessary for arbitrage to occur, and goods are diverted from markets in which the product is readily available (and usually lower priced) to the markets in which the product is in short supply (Banerji 1990).

The potential for some distributors or unauthorized third parties to free ride off full-service distributors' market development efforts (Mathewson and Winter 1984; Rubin 1990; Telser 1960) also facilitates the development of a gray market (Cross, Stephan, and Benjamin 1990). Although services such as advertising, pre- and postsales service, and warranty servicing (Coughlan et al. 2001) no doubt add value to the product, the revenue streams accruing to authorized distributors from their marketing efforts are vulnerable to appropriation (Anderson and Weitz 1986). Customer heterogeneity on services also is likely to foster gray markets. Some customers may value and even require services, whereas others may not be willing to pay for perceived "frills." Enterprising diverters would be quick to provide the actual product with little or no augmentation to those unwilling to pay the high price for the "bundled" product or service offering, thus realizing incremental (albeit unauthorized) sales to an additional, hitherto ignored customer segment (Ahmadi and Yang 2000; Yang, Ahmadi, and Monroe 1998). The preceding discussion leads to the following hypothesis:

H_5: The greater the (a) price differential, (b) premium positioning, (c) product scarcity, (d) free-riding potential, and (e) customer heterogeneity on services in a market, the greater is the likelihood of gray market incidence.

Research Method

Study 1: Field Survey of Gray Marketing of Personal Care Products

Our sampling frame comprised U.S.-based manufacturers of branded personal care products (e.g., skin toners, hair and nail care products, medicated soap) that distribute through wholesaler distributors in the professional salon industry. The reliance on external (non-company-owned) distribution intermediaries and the high value-to-weight ratio of the products in this study make this a particularly appropriate context in which to test our hypotheses. According to the Beauty and Barber Supply Institute, the association representing personal care product manufacturers and distributors, more than half of its members had experienced gray marketing of their products in some form or other. Thus, the phenomenon of gray markets is well known to most industry participants.

We defined the term "gray marketing" for informants, and then we asked them to focus on a geographic market in which they had experienced, or were likely to experience, gray marketing. After they identified such a market, we asked them to report their enforcement behavior (i.e., severity, certainty, and speed) in general rather than in response to a particular violation. We did this to assess deterrence attributable to a general propensity for enforcement rather than a particular response. Consistent with previous channel-related research in marketing (Anderson, Hakansson, and Johanson 1994; Anderson and Weitz 1992; Antia and Frazier 2001; Cavusgil and Sikora 1988; Dutta, Heide, and Bergen 1999; Dwyer, Schurr, and Oh 1987), we solicited manufacturers' evaluations of their own severity, detection ability, and speed of enforcement behavior. We developed and refined the survey instrument on the basis of a literature review, interviews with channel management personnel from the personal care products industry, and a mailed pretest.

Measures. Our dependent variable of interest, incidence (INCID), is a dichotomy that describes whether gray marketing occurred during the previous two years. We coded INCID as 1 if the informants indicated that they were aware of diversion of any product from their product line within the specified time period and 0 if otherwise.

We measured the independent variables with multi-item summated scales. The final item sets, response formats, individual item loadings, composite reliability, and average variance explained (AVE) for each scale appear in Appendix A. The preceding metrics provide evidence that the psychometric properties of our measures are satisfactory. Table 1 displays the correlation matrix and descriptive statistics for the variable set. In the following sections, we describe the measurement approach for each variable.

The manufacturer's enforcement severity (SEV) consists of a four-item Likert scale. We derived the items that constitute the scale from the work of Antia and Frazier (2001), but we modified them to ground the measure in our research context. Detection ability (DETECT) reflects the extent to which the manufacturer is able to evaluate the extent and nature of instances of product diversion. We derived the scale for this construct from the work of Dutta, Heide, and Bergen (1999) and, again, adapted it to our research context. The final scale is composed of four reverse-coded Likert items. Enforcement speed (SPEED) refers to the time elapsed between detection of the violation and the manufacturer's undertaking of corrective action. We developed the four specific items that constitute the scale using prior research in sociology and social psychology (Erickson, Gibbs, and Jensen 1977; Howe and Brandau 1988; Howe, Brandau, and Loftus 1996) and modified them on the basis of field interviews.

Price differential (PRDIFF) refers to the extent to which the manufacturer charges different prices for the product across markets. Our scale for this variable consists of a three-item Likert measure. Premium positioning (PREM) describes the extent to which the manufacturer positions the specific product as a high-end "premium" brand. On the basis of our literature search and field interviews, we developed a four-item Likert scale for this variable. Product scarcity (SCARCE) reflects the manufacturer's inability to supply the relevant market with enough product to satisfy demand. We measured the extent of product scarcity with a four-item Likert scale.

Potential for free riding (FR) reflects the extent to which the manufacturer believes that other authorized distributors or unauthorized third parties can gain from the sales efforts of its authorized distributors in that market. During the interviews we conducted, it became clear that we needed to capture the risk of both types of free riding. This corresponds to a second-order confirmatory factor model, in which the observed six items (three items reflecting each first-order factor) are hypothesized to originate from the two first-order factors (AUTHFR and UNAUTHFR), and in turn, the first-order factors originate from a second-order factor.

Customer heterogeneity (HETER) reflects the extent to which the specific market is characterized by distinct customer segments, in which each segment values services differently. We measured this variable with a three-item Likert scale, which we developed and refined on the basis of pretests.

Data collection. The Beauty and Barber Supply Institute agreed to send out a letter on our behalf that explained the purpose of the study and informed each of its members to expect further communication from us. A week later, we contacted each firm by telephone to introduce ourselves, ascertain the willingness to participate, and identify suitably qualified key informants. The modal designation of the informants was general manager; sales and channels managers were also well represented. Of the 581 firms called, we were unable to contact 21, 26 expressed their unwillingness or inability to participate, and another 41 had open lines or used company-owned distribution channels, leaving us with a final sample of 493.

We then sent out a survey packet by mail to the identified key informants at the remaining 493 manufacturer firms. We dropped an additional 20 informants from the study because they indicated that the study did not apply to their current business context. Two further rounds of callbacks and two subsequent mailings yielded a total of 104 usable questionnaires, representing a 22% response rate. Of the 104 responding firms, 40 reported experiencing at least one incidence of gray marketing, and 64 reported otherwise. Nonsignificant differences between early and late respondents (Armstrong and Overton 1977) on the focal variables of the study suggest that nonresponse bias is unlikely. Post hoc checks indicate acceptable levels of involvement ($M = 3.92$ on a five-point scale; $SD = 1.37$) and knowledge ($M = 4.34$; $SD = .97$) with respect to their firms' dealings in the specific product-market.

Measure validation. Given our sample size and the number of constructs, we conducted confirmatory factor analyses (CFAs) on groups of maximally similar constructs (see Moorman and Miner 1997), namely, the enforcement characteristics (severity, speed, and detection ability) and the contextual variables (price differences, customer heterogeneity, premium positioning, and supply shortage). The higher-order construct conceptualization of free-riding potential necessitated the estimation of 14 additional parameters, thus requiring us to conduct a separate CFA for this construct to meet sample size requirements. The CFA model diagnostics (see Table 2) suggest unidimensionality of the reflective scales. All items loaded on their prespecified constructs and had t-values significant at .05, providing evidence of convergent validity. Discriminant validity of the scales is further supported by the Lagrange-multiplier tests; none of the possible cross-loadings exceeds the critical value of the chi-square with one degree of freedom (Speier and Venkatesh 2002).

Given the significant correlation between enforcement severity and its speed, we undertook a rigorous test of the discriminant validity of these constructs in particular. We asked three academics with expertise in the area of enforcement to sort the eight items that purported to measure severity and speed. All three experts were able to perform the item-sorting task with full accuracy. Interitem correlations among the items measuring severity (average $\rho_{severity} = .68$, $p < .0001$) and speed (average $\rho_{speed} = .73$, $p < .0001$) are at a significantly higher level than cross-construct item correlations (average $\rho_{speed,severity} = .53$, $p < .0001$); t-tests of differences further suggest that the differences in correlations are significant. We also found that the AVE for each construct was greater than the squared correlation between the constructs (Fornell and Larcker 1981), thus providing further evidence of discriminant validity between severity and speed of enforcement. In all, the preceding tests suggest

TABLE 1
Study 1: Descriptive Statistics

	Correlation Matrix												
	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Incidence (INCID)	—												
2. Enforcement severity (SEV)	.08	.89											
3. Enforcement speed (SPEED)	.07	.68*	.92										
4. Detection ability (DETECT)	−.14	.28*	.29*	.85									
5. Price differential (PRDIFF)	.08	−.07	−.15	−.09	.79								
6. Free-riding potential (FR)	38*	.15	.11	−.31*	−.11	—							
7. Product scarcity (SCARCE)	.08	−.04	−.24*	−.01	−.02	.08	.87						
8. Premium positioning (PREM)	.15	.13	.11	−.00	−.01	.26*	−.02	.89					
9. Customer heterogeneity (HETER)	.26*	.26*	.20*	−.02	−.06	.22*	−.10	.12	.77				
10. SEV × DETECT	−.28*	−.01	−.07	.23*	−.02	−.28*	−.02	−.05	−.04	—			
11. SEV × SPEED	−.10	−.33*	−.47*	−.05	−.09	−.14	.09	.07	−.09	.31*	—		
12. SPEED × DETECT	−.19	−.07	−.10	.31*	−.06	−.23*	.01	.08	−.05	.80*	.27*	—	
13. SEV × SPEED × DETECT	−.12	.38*	.36*	.59*	−.06	−.14	.09	.11	−.11	−.03	−.12	−.07	—
Number of items	1	4	4	4	3	6	4	4	3	—	—	—	—
Minimum	0	4	4	4	3	6	4	4	3	—	—	—	—
Maximum	1	20	20	20	14	30	20	20	15	—	—	—	—
M	.38	13.46	13.57	9.67	5.78	20.45	7.66	14.27	9.34	.68	.28	.29	−.07
SD	.49	4.27	4.31	4.14	2.63	5.02	3.37	4.24	2.7	1.29	1.02	1.02	1.54

*r is significant at $p < .05$.

Notes: Diagonal entries represent alpha reliability values, where applicable.

TABLE 2
Study 1: CFA Results for Reflective Measures

Construct Groups	χ^2	*p* Value	Average Off-Diagonal Squared Residual	Comparative Fit Index	Incremental Fit Index	Tucker–Lewis Index	Lowest t-Value of Loading
SEV, SPEED, DETECT	$\chi^2_{51} = 90.03$	.001	07	.96	.96	.94	5.81
FR	$\chi^2_{8} = 16.61$	.03	.10	.98	.98	.96	7.36
PRDIFF, SCARCE, PREMIUM, HETER	$\chi^2_{71} = 89.9$	.07	.05	.97	.97	.96	5 26

that severity and speed of enforcement have adequate discriminant validity.

Model specification. We test our hypotheses using a maximum likelihood estimation–based logistic regression model, first specifying the main effects only (Equation 1) and then adding the two- and three-way interactions among the enforcement characteristics (Equation 2) as follows:

$$\text{(1)} \quad P(\text{INCID}_i = 1) = \frac{\exp\left(\beta_0 + \sum_{1}^{8} \Sigma\beta_j X_{ij}\right)}{1 + \exp\left(\beta_0 + \sum_{1}^{8} \Sigma\beta_j X_{ij}\right)}, \text{ and}$$

$$\text{(2)} \quad P(\text{INCID}_i = 1) = \frac{\exp\left(\beta_0 + \sum_{1}^{12} \Sigma\beta_j X_{ij}\right)}{1 + \exp\left(\beta_0 + \sum_{1}^{12} \Sigma\beta_j X_{ij}\right)},$$

where INCIDi = 1 if firm i reports at least one gray market violation and 0 if otherwise, and

X_{i1} = SEV,
X_{i2} = DETECT,
X_{i3} = SPEED,
X_{i4} = PRDIFF,
X_{i5} = FR,
Xi_6 = SCARCE,
Xi_7 = PREM,
Xi_8 = HETER,
X_{i9} = SEV × DETECT,
X_{i10} = SEV × SPEED,
X_{i11} = SPEED × DETECT, and
X_{i12} = SEV × SPEED × DETECT.

To reduce the potential multicollinearity arising from multiplicative interaction terms (Aiken and West 1991) and to yield coefficient estimates within the observable range of the independent variables (Friedrich 1982; Jaccard, Turrisi, and Wan1990), we mean centered all the explanatory variables and interactions created as the product of the mean-centered components (see Brown, Dev, and Lee 2000; Jaccard, Turrisi, and Wan 1990; Rokkan, Heide, and Wathne 2003).[1]

Table 3 displays the estimation results for both equations. The likelihood ratio for the main effects vector of coefficients is not significantly different from zero ($\chi^2_{(3)}$ = 1.58, p = .66). In contrast, the likelihood ratio is significantly different from zero for the two-way ($\chi^2_{(3)} = 6.69$, p = .08) and three-way ($\chi^2_{(1)} = 2.82$, p = .09) interactions. Furthermore, the interaction model (χ^2 = 35.71, p = .0004) gives us confidence that the null hypothesis of all the coefficients being zero can be rejected. The model correctly classifies 80% of the observations, suggesting that the inclusion of the interaction terms is warranted. Therefore, we discuss the findings of this study as inferred from the estimation of the interaction model.[2]

Study 1 results. None of the three enforcement characteristics (severity, detection ability, and speed) has a significant impact on deterrence (incidence) in isolation (b_1 = .11, b_2 = –.04, b_4 = .01). However, our examination of the full model coefficients requires us to separate the main effects from the interaction terms (for a full explanation, see Aiken and West 1991, p. 38; for a recent application in a channels context, see Brown, Dev, and Lee 2000). We do this by differentiating Equation 2 with respect to each of the three enforcement characteristics. Equation 3 shows how INCID varies as a function of SEV (for derivation and details on obtaining simple slopes, standard errors, and t-values, see Appendix B; for the results, see Table 4).

$$\text{(3)} \quad \frac{\partial \text{INCID}}{\partial \text{SEV}} = b_2 + b_3\text{DETECT} + b_5\text{SPEED} + b_7\text{SPEED} \times \text{DETECT}$$

---

[1] Main effects in this case represent the conditional effects of each enforcement characteristic on incidence at moderate (mean) levels of the other two facets. This interpretation is distinct from the "constant effect of one variable over all values of another variable" (Aiken and West 1991, p. 38) interpretation, as is the case with uncentered (raw) data. However, both approaches are consistent in that estimates of the latter may be recovered as a special case of the mean-centered approach, if required (Freidrich 1982).

[2] Table 1 shows evidence of considerable multicolinearity between the interaction variables and their components, such multicolinearity is known to inflate the standard errors of the parameter estimates, though the estimates themselves remain unbiased (Jaccard, Turrisi, and Wan 1990). The statistical significance of the higher-order interactions in the presence of the lower-order terms indicates that the imprecision due to multicolinearity does not pose a threat to validity (Buvik and John 2000). Parameter estimates and significance levels remain robust across both model specifications (the main effects–only and the fully specified interaction model), thus lending further confidence in our findings.

TABLE 3
Study 1: Parameter Estimates

Independent Variable	Main Effects Only		Full Model		
	Coefficient	t-Value	Coefficient (Interactions Included)	t-Value	Hypothesis Supported?[a]
Constant	– 661	–2.73***	–.519	–1.8**	
DETECT (b_1)	–.012	–.19	.109	1.20	H_1 (No)
SEV (b_2)	–.042	–.53	–.041	– 49	
SEV × DETECT (b_3)			–.055	–2.04**	H_2 (Yes)
SPEED (b_4)	.065	81	.010	.11	
SEV × SPEED (b_5)			–.009	–.59	H_3 (No)
SPEED × DETECT (b_6)			–.008	–.30	
SEV × SPEED × DETECT (b_7)			–.007	–1.68**	H_4 (Yes)
PRDIFF (b_8)	.147	1.60*	.145	1.51*	H_{5a} (Yes)
FR (b_9)	190	3.01***	206	2.89***	H_{5b} (Yes)
PREM (b_{10})	.035	.58	.056	.84	H_{5c} (No)
SCARCE (b_{11})	.094	1.31*	.121	1 45*	H_{5d} (Yes)
HETER (b_{12})	.201	2.08**	.194	1.82**	H_{5e} (Yes)
	$\chi^2_8 = 25.44$; $p = .0013$		$\chi^2_{12} = 35.71$; $p = .0004$		

*$p < .10$ (one-tailed test).
**$p < .05$ (one-tailed test)
***$p < .01$ (one-tailed test)
[a]Based on Interaction model estimates.

TABLE 4
Study 1: Impact of Enforcement Characteristics on Incidence

Enforcement Characteristic	Estimated Impact on Incidence (Simple Slope)	SE	t-Value
Main Effects			
SEV	.098	.145	.67
DETECT	246	.152	1.61
SPEED	–.046	.127	–.36
Interaction Effects			
Impact of SEV on INCID at various levels of DETECT (H_2)			
DETECT$_{low}$	.187	.137	1.35
DETECT$_{moderate}$	–.041	.084	–.49
DETECT$_{high}$	–.269	.137	–1.96*
Impact of SEV on INCID at various levels of SPEED (H_3)			
SPEED$_{low}$	–.028	127	–.224
SPEED$_{moderate}$	.010	.091	.110
SPEED$_{high}$	.048	.087	.556
Impact of SEV on Incidence at various levels of DETECT and SPEED (H_4)			
DETECT$_{high}$ and SPEED$_{high}$	–.43	.157	–2.85*
DETECT$_{high}$ and SPEED$_{low}$	–.11	.157	–.70
DETECT$_{low}$ and SPEED$_{high}$	.27	.157	1.80
DETECT$_{low}$ and SPEED$_{low}$	.10	.157	.65

*Significant at the .05 level (two-tailed test).

The t-value of .098 is not significant at the .05 level. Therefore, we are unable to reject the null hypothesis for SEV, which in turn enables us to reject the main effect we hypothesized in H_1. We evaluate the main effects of DETECT (∂INCID/∂DETECT = .246) and SPEED (∂INCID/∂SPEED = –.046) similarly (for estimates, see Table 4); in each case, the main effect is not significantly different from zero. We now proceed to examine the interaction coefficients.

We find that just one two-way interaction is statistically significant: SEV × DETECT ($b_3 = -.055$, $p < .05$). On further probing (see Table 4), we find that strict enforcement behavior has a significant inverse relationship to gray market incidence when it is combined with high detection ability ($\partial INCID/\partial SEV = -.269$, $p < .05$). Note that no such deterrent impact occurs under low ($\partial INCID/\partial SEV = .187$, $p > .10$) or even moderate ($\partial INCID/\partial SEV = -.041, p > .10$) levels of detection ability. The preceding result underscores the importance of matched and high levels of severity and detection ability, in support of the interaction we hypothesized in H_2. We fail to find support for the exploratory hypotheses with respect to SPEED's pairwise interaction with SEV ($b_5 = -.009$). That is, SPEED does not seem to confer any deterrent impact in combination with SEV or with DETECT.

In Table 3, we find that the three-way interaction among SEV, SPEED, and DETECT is statistically significant ($b_7 = -.007, p < .05$), in support of H_4, suggesting that the combination of all three facets of enforcement has a deterrent impact. Post hoc probing of this significant interaction reveals that increasing SEV results in deterrence only when SPEED and DETECT both have high levels (the simple slope of SEV on INCID is negative and significant). As can be seen in Table 4, if either SPEED or DETECT is low, the simple slope of SEV on INCID is not significant.

The significant and inverse relationship between SEV × DETECT and INCID suggests that the combination of high levels of SEV and DETECT is sufficient to exert a deterrent impact on gray market activity. Furthermore, the significant three-way interaction suggests that when SPEED is added to the combination of SEV and DETECT, deterrence is enhanced. Our inability to find similar effects for the pairwise combinations involving SPEED or for a main effect of SPEED suggests that speed must be accompanied by severe action and high detection ability to achieve incremental deterrence. Swift enforcement provides an additional deterrent "boost," but not by itself.

Our hypothesis with respect to the gray market–specific determinants of incidence finds significant support for four of the five contextual variables. The presence of price differentials among markets has a significant, though marginal, effect on gray market incidence ($b_8 = .145, p < .10$). Free-riding potential has a strong, significant, direct relationship to gray market activity ($b_9 = .206, p < .01$). We do not find any support for the hypothesized relationship between premium positioning and gray market incidence ($b_{10} = .056$). However, we find that product scarcity is positively related to the likelihood of gray market violations, though marginally so ($b_{11} = .121, p < .10$). We also find a strong positive relationship between customer heterogeneity on services and gray market incidence ($b_{12} = .194, p < .05$).

To summarize, our examination of the main and interactions effects suggests that none of the characteristics of enforcement has any significant deterrent impact in isolation. Matched and high levels of detection ability and severity are prerequisites for achieving deterrence, and swift enforcement boosts the preceding interaction.

Study 2: Experiment

We conducted Study 2 to replicate and extend the results of Study 1 with an alternative methodology, context, and measures. First, we conducted Study 2 as an experiment because the dimensions of enforcement severity, detection ability, and speed are likely to be naturally associated in field contexts. Manufacturers that are concerned with gray market activities tend to design detection systems that enable them to identify problems accurately and then react both quickly and severely against the offending firm. In contrast, manufacturers that are not affected by gray market activities are likely to be relatively lax on all three enforcement facets. Given this logic, some combinations of enforcement strategies are unlikely to occur naturally in field settings, such as high severity but low speed and certainty of detection or low severity and low speed but high certainty. The coincident use of these enforcement characteristics is evident in our first study with significant correlations between severity and speed ($r = .68, p < .05$), severity and detection ($r = .28, p < .05$), and detection and speed ($r = .29$, $p < .05$). Therefore, Study 2 uses an experimental design to examine the full range of possible enforcement strategies and to use orthogonal manipulations that enable us to assess the independent effects of severity, detection, and speed on deterrence.

Second, Study 2 assesses the robustness of the results of Study 1 by examining deterrence from the perspective of the dealer rather than the manufacturer. Given that it is dealers' evaluations of the effectiveness of the enforcement facets that affect their willingness to engage in gray market activities, it is important to understand their perceptions. Study 2 also differs from Study 1 in its use of a continuous rather than dichotomous measure of deterrence, thus enabling us to estimate the psychometric properties of the dependent variable.

Procedure. Participants were 112 MBA students who were recruited by student representatives of an on-campus exchange program at a major university. A donation of $10 was made toward the program on behalf of each participant who completed the study. Sixty-four percent of the participants were male, their average age was 29 years, and they had an average of three years of management experience. Seventy-seven percent of participants had at least one year of experience as "a supplier or dealing with suppliers in a business situation."

After a brief introduction to a study on manufacturer–distributor relationships, participants were randomly assigned to one condition in a 2 (severity: high versus low) × 2 (certainty: high versus low) × 2 (speed: fast versus slow) full-factorial between-subjects design. Participants read a scenario that explained that an electronics components dealer had violated its agreement with a manufacturer by selling $100,000 of integrated circuits outside its territory. The scenario included manipulations of the severity, certainty, and speed of the manufacturer's response. The method assumes that participants project themselves into the hypothetical situation and provide answers that reflect how dealers would actually respond to the situation outlined in the scenario. Evidence suggests that projective methods can accurately represent participants' underlying attitudes and behaviors (e.g., Fisher 1993) and that the judgments of individual managers can provide important insights into organization-level strategies (e.g., Chandy, Prabhu, and Antia 2003). The scenario and manipulations appear in Appendix C. Participants then responded to a series of questions.

The dependent variable was a three-item measure that reflected the degree to which the offending dealer would be deterred from engaging in future gray market activities by the manufacturer's enforcement activities. The items included the following: "In future, the offending dealer would definitely comply with the manufacturer's regulations governing unauthorized sales," "In future, the offending dealer would not dare [to] disregard the resale restrictions of the manufacturer," and "In future, the offending dealer would be deterred from selling the manufacturer's products in an unauthorized manner." The items were measured on a five-point Likert scale, anchored by "very unlikely" and "very likely." The mean of the summated scale items was 2.74, with a standard deviation of .97 and a coefficient alpha of .86.

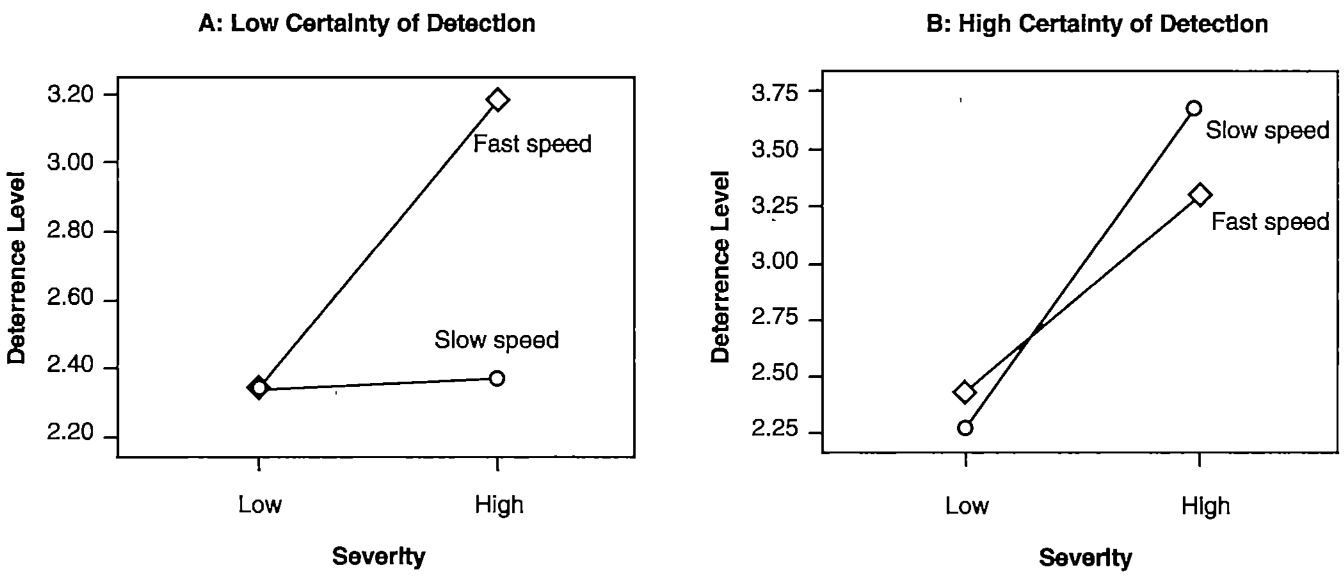
Analysis. We used a full-factorial analysis of variance to examine the manipulation checks and to test the hypotheses. We found a significant effect of the detection manipulation on the item "The manufacturer is very likely to catch dealers who undertake gray market activities" (F(1, 111) = 131.51, $p < .01$), a significant effect of the severity manipulation on the item "The manufacturer's response to gray market transactions is severe" (F(1, 111) = 76.10, $p < .01$), and a significant effect of the speed manipulation on the item "The manufacturer punishes gray market activity immediately" (F(1, 111) = 122.77, $p < .01$). The manipulations had no unintended main or interaction effects.

In terms of the hypotheses, we found significant main effects of certainty (F(1, 111) = 5.20, $p < .05$) and severity (F(1, 111) = 22.81, $p < .01$) on deterrence. In terms of interactions, the two-way interaction between certainty and severity was significant (F(1, 111) = 5.23, $p < .05$), as was the three-way interaction effect (F(1, 109) = 4.17, $p < .05$). Examination of the simple effects reveals that severity has a positive effect on deterrence in three of four possible instances: high certainty/high speed (F(1, 27) = 10.32; $M_{\text{low severity}} = 2.43 > M_{\text{high severity}} = 3.29$, $p < .01$), low certainty/high speed (F(1, 27) = 3.48; $M_{\text{low severity}} = 2.38 > M_{\text{high severity}} = 3.15$, $p < .10$), and high certainty/low speed (F(1, 27) = 19.97; $M_{\text{low severity}} = 2.31 > M_{\text{high severity}} = 3.62$, $p < .001$). As we predicted, severity has no effect on deterrence when both certainty and speed are low ($p > .90$). Note that when certainty and severity are high, speed has no effect on deterrence ($p > .30$). We included participants' gender and years of supervisory experience as covariates, but they were not significant ($p > .20$). The analysis of variance interaction results appear in Figure 1.

Study 2 results. The results are consistent with those of Study 1 in three of four possible effects of severity on deterrence. Consistent with Study 1, the results of Study 2 support the view that severity of response alone is insufficient to reduce gray market activities. Increasing the severity of the punishment for gray market violations had no effect when dealer violations were unlikely to be detected and when violators could expect a long period of time before they would be punished. It seems logical that the threat of even the most severe punishment will be ineffective if it has a low likelihood of imposition and if it will only take place in the distant future.

Increasing the severity of the manufacturer's response to gray market activity was effective when it was used in concert with systems that ensured a high likelihood of detection, immediate responses to violations, or both. Consistent with Study 1, participants indicated that future gray market activities would be reduced when the likelihood of detection was high, regardless of whether enforcement speed was fast or slow. The results support the view that deterrence is a function of both the severity and the likelihood of punish-

FIGURE 1
Study 2: Certainty × Speed × Severity Interaction Effect



ment. In contrast to the Study 1, Study 2 found that severity had a deterrence effect when speed was high but the likelihood of detection was low. By implication, speed appears to compensate for detection certainty in some conditions. Despite a low probability of detection and, thus, punishment, participants believed that dealers would be less likely to reoffend if severe punishments were meted out quickly.

Discussion

The results of the two studies are consistent in their support of the deterrence doctrine, despite varying dramatically on various dimensions, including methodology (field survey versus experiment), perspective (manufacturer versus distributor), and measurement of the dependent variable (dichotomous measure of actual incidence versus multi-item measure of future recidivism). Both studies indicate that the severity, certainty, and speed of enforcement interact in their effects on deterrence. The results have important implications for both research and practice in the area.

Research Implications

Prior research in marketing has implicitly taken the relationship between enforcement severity and deterrence to be an article of faith. So well entrenched is this assumption that there has been no attempt to test the efficacy of enforcement in a marketing context. Our central finding is that, by itself, enforcement severity does not deter gray market incidence. As a result, an expanded conceptualization of enforcement is necessary. Our study also sheds light on how the facets of enforcement work together to minimize gray market violations. Across both studies, none of the three characteristics of enforcement (i.e., severity, certainty, and speed) has deterrent value alone. Rather, deterrence is most likely to occur when the penalties for gray market violations are severe, when manufacturers are able to detect violations or mete out punishments in a timely fashion, or both.

The introduction of speed of response further extends knowledge of the subtleties of enforcement. Specifically, we find that swift enforcement bolsters the deterrent impact of high levels of detection ability and severity. The ability to increase the effectiveness of the latter combination is particularly interesting considering that, by itself, enforcement speed has little to contribute to deterrence objectives. This finding is of great importance to the marketing literature on enforcement; it emphasizes the dual imperative for an expanded conceptualization of enforcement, and it forces us to examine the deterrent impact of enforcement rather than take it for granted as research has done to date.

In examining the deterrent impact of enforcement, this study also links the evolving marketing literature on opportunism (Brown, Dev, and Lee 2000; Dahlstrom and Nygaard 1999; John 1984; Wathne and Heide 2000) and enforcement (Antia and Frazier 2001; Bergen, Heide, and Dutta 1998; Dutta, Bergen, and John 1994). The preceding literature bases have tended to develop along distinct paths, despite the obvious linkages. To the best of our knowledge, ours is the first study to test the assumed deterrent efficacy of enforcement. Our findings set the stage for a synthesis of these still disparate research streams.

Our choice of empirical context facilitates an additional important contribution. To date, the literature on gray markets has tended to develop along two lines: analytical models (Ahmadi and Yang 2000; Banerji 1990; Bucklin 1993; Coughlan and Soberman 1998) and descriptive case studies (Antia and Everatt 2000; Cespedes, Corey, and Rangan 1988). By their very nature and design, the preceding efforts emphasize deeper understanding of a select few issues relevant to gray markets. What is needed is an integration and subsequent large-sample validation of the commonly suggested drivers of gray market activity. We address this crucial gap in the understanding of gray markets by developing an integrative model of gray market incidence and by validating this model with both field data and an experiment. Our study represents an attempt to "take stock" of received thought on gray market incidence.

Managerial Implications

Until now, the vocabulary and emphasis of the business and academic press have been limited to enforcement severity. Yet our results suggest that current prescriptions calling for stricter measures against gray market participants are likely to be ineffective. Managers who want to make progress against gray market incursions must invest in systems that increase the ability to both detect and quickly punish violators. It is only when severe enforcement behavior is combined with certainty and/or speed that gray market incidence is significantly curbed.

Our work points out the need to use an integrated approach to the management of distribution channel relations. In addition to devoting attention to enforcement, manufacturers should provide adequate incentives to authorized distributors for the continued provision of valued services in the right mix to the right targeted segment. Due consideration should also be given to pricing and supply issues, in acknowledgment of their potential impact on gray market activity. The combination of all the preceding measures is more likely to lead to desired deterrence objectives.

Limitations and Future Research Directions

Although the two studies we reported herein provide several contributions to academics and managers alike, much work remains to be done. The term "gray marketing" encapsulates a variety of unauthorized sales, including cross-border parallel importation, diversion, and domestic bootlegging, to name but a few. Although cognizant of these subtle distinctions, we were unable to control for the location and specific type of gray market activity because of nonresponse on the particular items in our survey. We have taken initial steps toward understanding the degree to which the deterrence doctrine is generalizable by testing it from both the manufacturer's (Study 1) and the dealer's (Study 2) perspective, but a variety of other industry and violation contexts remain unexamined.

Our empirical context imposed certain constraints on our construct operationalizations. First, the absence of archival data on gray market incidence and the steps taken by manufacturers in response forced us to restate our conceptual framework in terms of detection ability rather than the notion of certainty, as originally conceptualized by the deterrence doctrine. Second, in Study 1, we measured gray market incidence as a binary response in an attempt to strike a balance between examining gray markets in a meaningful manner and minimizing the potential for nonresponse attributable to what could be perceived as excessive probing of a "hot-button" channel issue. In Study 2, we were able to replicate the findings of Study 1 using a continuous measure of deterrence in an experimental context. We consider our studies a stimulus for more sophisticated attempts to capture the intricacies of channel violations and manufacturers' subsequent responses.

Consistent with a well-established body of channels research, we asked manufacturer informants to report their own behavior. In Study 2, we considered the dealer's perspective and found similar results. It would be useful for research to incorporate the simultaneous views of both manufacturers and dealers to understand this issue further.

APPENDIX A
Measures

Enforcement Severity (Composite Reliability = .90; AVE = .69)

We have a policy of full enforcement of our sales agreements.	.75
We are well known for our strict policy on product diversion.	75
We have a tough stance on product diversion.	.95
We (would) take strict action against unauthorized sales.	.85

Enforcement Speed (Composite Reliability = .92; AVE = .73)

Our response to violations is (would be) instantaneous.	.88
We (would) take immediate action against violations.	.83
Very little time (would) elapse between detection of violations and our response to them	.87
Our enforcement response process is (would be) very timely.	.85

Detection Ability (Composite Reliability = .87; AVE = .63)

At a given time, it would be difficult to evaluate the extent of product diversion. (R)	.82
Determining compliance with resale restraints requires a great amount of effort on our part. (R)	.56
It would be difficult for us to evaluate the extent to which our product is diverted. (R)	.95
Our evaluation of the extent of unauthorized sales is based on very "fuzzy" information. (R)	78

Price Differential (Composite Reliability = .80; AVE = .58)

We maintain a uniform pricing policy between markets. (R)	.88
In general, we try to keep price differences between markets to a minimum. (R)	.80
The price we charge for the product in each market varies considerably.	.57

Free-Riding Potential

Unauthorized Free Riding (Composite Reliability = .85; AVE = .65)

Customers may learn about our product from our distributors and purchase it from unauthorized sources.	.91
Our distributors' presales services might stimulate unauthorized sales of the product.	.80
Unauthorized distributors could benefit from the market development efforts of our authorized distributors.	.70

Authorized Distributor Free Riding (Composite Reliability = .92; AVE = .80)

Our authorized distributors' sales efforts increase sales of other authorized distributors of this product.	.84
The services provided by one authorized distributor may help other authorized distributors' sales.	.92
Authorized distributors could benefit from the market development efforts of other authorized distributors.	.92

Product Scarcity (Composite Reliability = .88; AVE = .64)

We have trouble keeping up with demand for this product.	.61
Order fulfillment for this product is frequently delayed due to production constraints.	.82
We frequently experience stock-outs of this product.	.93
We do not produce enough product to satisfy demand.	.82

Premium Positioning (Composite Reliability = .89; AVE = .67)

Our brand name commands a significant price premium in this market.	.74
Our company is considered a market leader in this product category.	.86
We are able to leverage our strong brand name to a great extent in this market	.87
Our products have a strong reputation in this market	.80

Customer Heterogeneity (Composite Reliability = .78; AVE = .54)

Customers in this market require very different service levels.	.61
Our customers have very dissimilar product preferences.	.71
Customers in this market differ considerably in their preference for service.	.60

Notes: All items are measured on a five-point Likert scale, anchored by "strongly agree" and "strongly disagree." (R) = reverse-scored items.

Appendix B
Post Hoc Probing of Interactions

Two-Way Interaction Between SEV and DETECT

The objective of this post hoc probing is to ascertain the impact of severity on incidence for different levels of detection ability. Accordingly, we rearrange the terms of Equation 2 to obtain the simple regression of INCID on SEV at plus and minus one standard deviation of DETECT, respectively (Aiken and West 1991, p. 13). This yields two simple slope coefficients, indicating the impact of SEV on INCID at each level of DETECT.

From Table 3, Equation 2,

$$\text{INCID} = -.519 - .041\text{SEV} + .109\text{DETECT} - .055\text{SEV} \times \text{DETECT}.$$

If we rearrange the terms,

$$\text{INCID} = (-.041 - .055\text{DETECT})\text{SEV} + (-.519 + .109\text{DETECT}).$$

At $\text{DETECT}_L(\text{DETECT} = -4.1447)$,

$$\text{INCID} = (-.041 - .055 \times -4.1447)\text{SEV} + (-.519 + .109 \times -4.1447),$$

or

$$\text{INCID} = .187\text{SEV} - .971.$$

We obtain the simple slope coefficients similarly for each two-way interaction (see Table 4). Using the variance–covariance matrix of the beta estimators, we then obtain the standard error and t-value for each simple slope.

Three-Way Interaction Among SEV, SPEED, and DETECT

We rearrange the terms of Equation 2 yet again, this time replacing SPEED and DETECT with values plus and minus one standard deviation to represent high and low values of SPEED and DETECT, respectively.

From Table 2, Equation 2,

$$\begin{aligned}\text{INCID} = {} & -.519 + -.041\text{SEV} + .010\text{SPEED} + .109\text{DETECT} \\ & - .055\text{SEV} \times \text{DETECT} - .009\text{SEV} \times \text{SPEED} \\ & - .008\text{SPEED} \times \text{DETECT} \\ & - .007\text{SEV} \times \text{SPEED} \times \text{DETECT}.\end{aligned}$$

If we rearrange the terms,

$$\begin{aligned}\text{INCID} = {} & (-.041 - .055\text{DETECT} - .009\text{SPEED} \\ & - .007\text{SPEED} \times \text{DETECT})\text{SEV} \\ & + (-.519 + .010\text{SPEED} + 1095\text{DETECT} \\ & - .008\text{SPEED} \times \text{DETECT}).\end{aligned}$$

We then obtain the simple slopes of SEV for the four combinations of SPEED and DETECT:

At SPEED_H and DETECT_H ($\text{SPEED} = 4.3051$, $\text{DETECT} = 4.1447$), $\text{INCID} = -.355\text{SEV} - .167$;

At SPEED_L and DETECT_H ($\text{SPEED} = -4.3051$, $\text{DETECT} = 4.1447$), $\text{INCID} = -.183\text{SEV} + .032$;

At SPEED_H and DETECT_L ($\text{SPEED} = 4.3051$, $\text{DETECT} = -4.1447$), $\text{INCID} = .351\text{SEV} - .785$; and

At SPEED_L and DETECT_L ($\text{SPEED} = -4.3051$, $\text{DETECT} = -4.1447$), $\text{INCID} = .023\text{SEV} - 1.157$.

Using the variance–covariance matrix of the beta estimators, we obtain the standard error and t-value for each simple slope (see Table 4).

Appendix C
Study 2 Scenario and Experimental Manipulations

Avon is a company that manufactures integrated chips (ICs) for use in laptops and personal computers. Avon dealers are expressly prohibited from selling their ICs outside their territory (called gray market activity). Nevertheless, one of Avon's dealers recently sold a large order of Avon ICs (approximately $100,000 in sales) outside its authorized territory to get rid of excess inventory. Avon has the following procedures and policies related to gray market activities by its dealers:

Detection Manipulation

High: The dealer has a high probability of being caught. Avon is able to identify unauthorized sales *over 90% of the time* because of its sophisticated inventory management system.

Low: The dealer has a low probability of being caught. Avon is able to identify unauthorized sales *only 10% of the time* because of an unsophisticated inventory management system.

Severity Manipulation

High: The penalties are significant. If Avon identifies an unauthorized sale, it levies a fine equal to twice the value of the sale (in this case $200,000). If caught, dealers *lose a significant amount of money* on the gray market transaction.

Low: The penalties are not very significant. If it identifies an unauthorized sale, it levies a fine equal to 15% of the contract's value. If caught, dealers *still make a small profit* on the gray market transaction.

Speed Manipulation

Fast: Avon's inventory management system is able to detect gray market activities *immediately*. After a violation is identified, the dealer is notified that [it is] being fined within 24 hours. Consequently, dealers are notified and punished *before* they make any additional gray market sales.

Slow: Avon requires a dealer audit before it can confirm a suspected gray market activity, so contract violations are detected *six months after the fact*. After a violation is identified, the dealer is notified that [it is] being fined within 24 hours. Consequently, many dealers are notified and punished after they have made additional gray market sales.

REFERENCES

Ahmadi, Reza H. and B. Rachel Yang (2000), "Parallel Imports: Challenges from Unauthorized Distribution Channels," *Marketing Science*, 19 (Summer), 279–94.

Aiken, Leona and Stephen West (1991), *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage Publications.

Andenaes, Johannes (1971), "The Moral or Educative Influence of Criminal Law," *Journal of Social Issues*, 27 (2), 17–31.

Anderson, Erin and Barton A. Weitz (1986), "Make-or-Buy Decisions: Vertical Integration and Marketing Productivity," *Sloan Management Review*, 27 (Spring), 3–19.

——— and ——— (1992), "The Use of Pledges to Build and Sustain Commitment in Distribution Channels," *Journal of Marketing Research*, 24 (February), 18–34.

Anderson, James, Hakan Hakansson, and Jan Johanson (1994), "Dyadic Business Relationships Within a Business Network Context," *Journal of Marketing*, 58 (October), 1–16.

Antia, Kersi D. and Donna Everatt (2000), *Looks.com(A): A Grey Issue*. London, Ontario: Ivey Publishing.

——— and Gary L. Frazier (2001), "The Severity of Contract Enforcement in Interfirm Channel Relationships," *Journal of Marketing*, 65 (October), 67–81.

Antunes, George and A. Lee Hunt (1973), "The Impact of Certainty and Severity of Punishment on Levels of Crime in American States: An Extended Analysis," *Journal of Criminal Law and Criminology*, 64 (4), 486–93.

Armstrong, Scott J. and Terry S. Overton (1977), "Estimating Nonresponse Bias in Mail Surveys," *Journal of Marketing Research*, 14 (August), 396–402.

Assmus, Gert and Carsten Wiese (1995), "How to Address the Gray Market Threat Using Price Coordination," *Sloan Management Review*, 36 (Spring), 31–41.

Bailey, William C. and Ronald W. Smith (1972), "Punishment: Its Severity and Certainty," *Journal of Criminal Law, Criminology, and Police Science*, 63 (4), 530–39.

Ball, John C. (1955), "The Deterrence Concept in Criminology and Law," *Journal of Criminal Law, Criminology, and Political Science*, 46 (September–October), 347–54.

Banerji, Shumeet (1990), *A Theory of Gray Markets: The Case of the Personal Computer Industry*. Evanston, IL: Northwestern University.

Beccaria, Cesare (1963), *On Crimes and Punishments*. Indianapolis: Bobbs-Merrill.

Becker, Gary S. (1968), "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 76 (March–April), 169–217.

Bentham, Jeremy (1962), *The Works of Jeremy Bentham*. New York: Russell and Russell.

Bergen, Mark, Jan Heide, and Shantanu Dutta (1998), "Managing Grey Markets Through Tolerance of Violations," *Managerial and Decision Economics*, 19 (May), 157–66.

Brown, James R., Chekitan S. Dev, and Dong-Jin Lee (2000), "Managing Marketing Channel Opportunism: The Efficacy of Alternative Governance Mechanisms," *Journal of Marketing*, 64 (April), 51–65.

Bucklin, Louis P. (1993), "Modeling the International Gray Market for Public Policy Decisions," *International Journal of Research in Marketing*, 10 (4), 387–405.

Buvik, Arnt and George John (2000), "When Does Vertical Coordination Improve Industrial Purchasing Relationships?" *Journal of Marketing*, 64 (October), 52–64.

Cavusgil, S. Tamer and E. Sikora (1988), "How Multinationals Can Cope with Gray Market Imports," *Columbia Journal of World Business*, 23 (4), 75–85.

Cespedes, Frank V., Raymond E. Corey, and Kasturi V. Rangan (1988), "Gray Markets: Causes and Cures," *Harvard Business Review*, 66 (4), 75–83.

Chandy, Rajesh K., Jaideep C. Prabhu, and Kersi D. Antia (2003), "What Will the Future Bring? Technology Expectations, Dominance, and Radical Product Innovation," *Journal of Marketing*, 67 (July), 1–18.

Corey, Raymond E., Frank V. Cespedes, and Kasturi V. Rangan (1989), *Going To Market*. Boston. Harvard Business School Press.

Coughlan, Ann T., Erin Anderson, Louis W. Stern, and Adel I. El-Ansary (2001), *Marketing Channels*. Upper Saddle River, NJ: Prentice Hall.

——— and David A. Soberman (1998), "When Is the Best Ship a Leaky One? Segmentation, Competition, and Gray Markets," working paper, INSEAD.

Cross, James, James Stephan, and Robert E. Benjamin (1990), "Gray Markets: A Legal Review and Public Perspectives," *Journal of Public Policy & Marketing*, 9, 183–94.

Dahlstrom, Robert and Arne Nygaard (1999), "An Empirical Investigation of Ex Post Transaction Costs in Franchised Distribution Channels," *Journal of Marketing Research*, 36 (May), 160–70.

Diver-Stamnes, Ann C. and R. Murray Thomas (1995), *Prevent, Repent, Reform, Revenge: A Study in Adolescent Moral Development*. Westport, CT: Greenwood Press.

Duhan, Dale F. and Jane Mary Sheffet (1988), "Gray Markets and the Legal Status of Parallel Importation," *Journal of Marketing*, 52 (July), 75–83.

Dutta, Shantanu, Mark Bergen, and George John (1994), "The Governance of Exclusive Territories When Dealers Can Bootleg," *Marketing Science*, 13 (Winter), 83–99.

———, Jan Heide, and Mark Bergen (1999), "Vertical Territorial Restrictions and Public Policy: Theories and Industry Evidence," *Journal of Marketing*, 63 (October), 121–34.

Dwyer, Robert F., Paul Schurr, and Sejo Oh (1987), "Developing Buyer–Seller Relationships," *Journal of Marketing*, 51 (April), 11–27.

Erickson, Maynard L., Jack P. Gibbs, and Gary F. Jensen (1977), "The Deterrence Doctrine and the Perceived Certainty of Legal Punishment," *American Sociological Review*, 42 (April), 305–317.

Fisher, Robert J. (1993), "Social Desirability Bias and the Validity of Indirect Questioning," *Journal of Consumer Research*, 20 (September), 303–315.

Fornell, Claes and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39–50.

Freidrich, R.J. (1982), "In Defense of Multiplicative Terms in Regression Equations," *American Journal of Political Science*, 26 (November), 797–833.

French, John R. and Bertram Raven, eds. (1959), *The Bases of Social Power*. Ann Arbor: University of Michigan Institute for Social Research.

Ghosh, Mrinal and George John (1999), "Governance Value Analysis and Marketing Strategy," *Journal of Marketing*, 63 (Special Issue), 131–45.

Gibbs, Jack P. (1968), "Crime, Punishment, and Deterrence," *Social Science Quarterly*, 48 (March), 515–30.

——— (1975), *Crime, Punishment, and Deterrence*. New York. Elsevier

Grasmick, Harold G. and Steven D. McLaughlin (1978), "Deterrence and Social Control," *American Sociological Review*, 43 (April), 272–78.

Gray, Louis N and J. David Martin (1969), "Punishment and Deterrence: Another Analysis," *Social Science Quarterly*, 50 (September), 389–95.

———, Alfred C. Miranne, David A. Ward, and Ben Menke (1982), "A Game Theoretical Analysis of the Components of Punishments," *Social Psychology Quarterly*, 45 (4), 206–212.

Hollinger, Richard C. and John P. Clark (1983), "Deterrence in the Workplace: Perceived Certainty, Perceived Severity, and Employee Theft," *Social Forces*, 62 (2), 398–418.

Howe, Edmond S. and Cynthia J. Brandau (1988), "Addictive Effects of Certainty, Severity, and Celerity of Punishment Judgments of Crime Deterrence Scale Value," *Journal of Applied Social Psychology*, 18 (9), 796–812.

———, ———, and Thomas C. Loftus (1996), "Integration of Certainty, Severity, and Celerity Information in Judged Deterrence Value: Further Evidence and Methodological Equivalence," *Journal of Applied Social Psychology*, 26 (3), 226–32.

Hufbauer, Clyde Gary, Jeffrey J. Schott, and Kimberly Ann Elliott (1990), *Economic Sanctions Reconsidered. History and Current Policy*, 2d ed. Washington, DC: Institute for International Economics.

Hwang, Suein L. (1999), "Tobacco: As Cigarette Prices Soar, a Gray Market Booms," *The Wall Street Journal*, (January 28), B1

Jaccard, James, Robert Turrisi, and Choi K. Wan (1990), *Interaction Effects in Multiple Regression*. Newbury Park, CA: Sage Publications.

John, George (1984), "An Empirical Investigation of Some Antecedents of Opportunism in a Marketing Channel," *Journal of Marketing Research*, 21 (August), 278–89.

Luckenbill, David F. (1982), "Compliance Under Threat of Severe Punishment," *Social Forces*, 60 (3), 811–25.

Manson, Allan (2001), *The Law of Sentencing*. Toronto Irwin Law.

Mathewson, George and Ralph Winter (1984), "An Economic Theory of Vertical Restraints," *Rand Journal of Economics*, 12 (Spring), 27–38.

Meier, Robert F. and Weldon T. Johnson (1977), "Deterrence as Social Control: The Legal and Extralegal Production of Conformity," *American Sociological Review*, 42 (April), 292–304.

Michael, James (1998), "A Supplemental Distribution Channel? The Case of U.S. Parallel Export Channels," *Multinational Business Review*, 6 (Spring), 24–35.

Moorman, Christine and Anne S. Miner (1997), "The Impact of Organizational Memory on New Product Performance and Creativity," *Journal of Marketing Research*, 34 (February), 91–107.

Myers, Matthew B. (1999), "Incidents of Gray Market Activity Among U.S. Exporters: Occurrences, Characteristics and Consequences," *Journal of International Business Studies*, 30 (1), 105–126.

——— and David A. Griffith (1999), "Strategies for Combating Gray Market Activity," *Business Horizons*, 42 (November–December), 2–8.

Nelson, Emily (1999), "On the Rack: Why Tommy Hilfiger Just Can't Seem to Stay Out of Local Wal-Mart," *The Wall Street Journal*, (January 28), A1.

Onkvisit, S. and J.J. Shaw (1989), *International Marketing: Analysis and Strategy*. Columbus, OH: Merrill Publishing Company.

Posner, Richard (1985), "An Economic Theory of the Criminal Law," *Columbia Law Review*, 85 (October), 1193–1231.

Rokkan, Aksel I., Jan B. Heide, and Kenneth H. Wathne (2003), "Specific Investments in Marketing Relationships: Expropriation and Bonding Effects," *Journal of Marketing Research*, 40 (May), 210–24.

Rubin, Paul (1990), *Managing Business Transactions*. New York: The Free Press.

Silberman, Mathew (1976), "Toward a Theory of Criminal Deterrence," *American Sociological Review*, 41 (June), 442–61.

Speier, Cheri and Viswanath Venkatesh (2002), "The Hidden Minefields in the Adoption of Sales Force Automation Technologies," *Journal of Marketing*, 66 (July), 98–111.

Stafford, Mark C., Louis N. Gray, Ben A. Menke, and David A. Ward (1982), "Modeling the Deterrent Effects of Punishment," *Social Psychology Quarterly*, 49 (December), 338–47.

Stigler, George (1970), "The Optimal Enforcement of Laws," *Journal of Political Economy*, 78 (May–June), 526–36.

Tedeschi, James T. (1976), *Social Psychology: Interdependence, Interaction and Influence*. New York: John Wiley & Sons.

Telser, Lester G. (1960), "Why Should Manufacturers Want Fair Trade?" *Journal of Law and Economics*, 3 (October), 86–105.

Wathne, Kenneth H. and Jan B. Heide (2000), "Opportunism in Interfirm Relationships: Forms, Outcomes, and Solutions," *Journal of Marketing*, 64 (October), 36–51.

Yang, B. Rachel, Reza H. Ahmadi, and Kent B. Monroe (1998), "Pricing in Separable Channels: The Case of Parallel Imports," *Journal of Product and Brand Management*, 7 (5), 433–40.

Velitchka D. Kaltcheva & Barton A. Weitz

When Should a Retailer Create an Exciting Store Environment?

Addressing inconsistent results in the literature, the conceptual framework that the authors develop proposes that the consumer's motivational orientation moderates the effect of the arousal produced by a store environment on the pleasantness of the environment. Two experiments support the framework. When consumers have a recreational motivational orientation, high arousal has a positive effect on pleasantness, but when consumers have a task-oriented motivational orientation, high arousal decreases pleasantness. In addition, high arousal increases consumer intentions to visit and make purchases in the store for recreationally oriented consumers, but it has a negative impact on shopping behavior for task-oriented consumers. Pleasantness mediates the effect of arousal on shopping behavior.

Recognizing the significant impact of store environment on consumer behavior, retailers devote considerable resources to store design and merchandise presentation activities. For example, Toys "R" Us spent $35 million to make its Times Square New York store "the ultimate toy store that is the personification of every kid's dream" (Prior 2001, p. 46). Some retailers (e.g., The Sharper Image, Bass Pro Shops, Barnes & Noble) attempt to create engaging environments in their stores that encourage shopping. In contrast, other retail chains (e.g., Costco, Home Depot) are successful with their minimalist warehouse-style shopping environments (Levy and Weitz 2004). With respect to electronic retailing contexts, it has been suggested that the most effective and desirable Web sites are those that are challenging to consumers and heighten their arousal level (Novak, Hoffman, and Yung 2000; Tedesco 2000). However, research has actually found no relationship, and even a negative relationship, between the use of arousal-inducing Web design elements and the pleasantness of the Internet experience (Da Cunha, Chakravarti, and Weitz 2003).

Although considerable research has examined the impact of various store environment elements on shopping behavior, the existing research does not provide an explanation for the differential effectiveness of arousal levels in store environments or managerially relevant guidelines for selecting the appropriate arousal level for a specific store environment. The objective of this article is to present and provide empirical support for a theoretical framework that outlines the impact of arousal-inducing features in a store environment on consumer affective responses to the environment and subsequent shopping behaviors in the environment. This theoretical framework proposes that an important situational variable—namely, the consumer's motivational orientation in the store—moderates the effect of arousal on pleasantness.

The next section briefly reviews research that investigates the impact of store environment on shopping behavior. This review highlights some inconsistent and, in some cases, contradictory findings, suggesting the need to consider moderator variables. Then, we propose a theoretical framework that focuses on the moderating effect of motivational orientation and resolves some of the inconsistent findings. We describe two laboratory experiments that test the theoretical framework. The article concludes with a discussion of the limitations of our research and its managerial implications.

Research on Environment and Shopping Behavior

Mehrabian and Russell's Conceptualization of the Environmental Impact on Behavior

Mehrabian and Russell's (1974) seminal conceptualization is the basis of most research on the impact of environmental factors on shopping behavior. Mehrabian and Russell propose that environment sensory variables, the environment's information rate (the amount of information in the environment), and individual differences in affective response influence people's affective responses to the environment, which in turn induce people to approach or avoid the environment. Mehrabian and Russell identify three dimensions of affective response: pleasantness, arousal, and dominance. However, subsequent research has found that pleasantness and arousal explain most of the variance in affect and behavior, and thus research attention has been focused on these two variables (Russell 1978; Russell and Pratt 1980).

Empirical Research on the Environmental Impact on Shopping Behavior

Table 1 summarizes the research investigating environmental effects on shopping behavior. The table shows that in almost all cases, pleasant shopping environments positively affect "approach" shopping-behavior variables, such as

---

Velitchka D. Kaltcheva is an assistant professor, Marketing Department, College of Business Administration, University of Miami (e-mail: vkaltche@exchange.sba.miami.edu). Barton A. Weitz is a professor and J.C. Penney Eminent Scholar, Warrington College of Business Administration, and Executive Director of the David F. Miller Center for Retailing Education and Research, University of Florida (e-mail: bart.weitz@cba.ufl.edu). The authors appreciate the comments and suggestions of Richard Lutz and Alan Sawyer.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 70 (January 2006), 107–118

TABLE 1
Empirical Research on the Impact of Pleasantness and Arousal on Shopping Behavior

Shopping Behaviors	Pleasantness	Arousal
Approach shopping behavior	P: Baker, Levy, and Grewal (1992) P: Donovan and Rossiter (1982) P: Hui and Bateson (1991) O: Spies, Hesse, and Loesch (1997) P: Sweeney and Wyber (2002)	P: Baker, Levy, and Grewal (1992)[c] P: Donovan and Rossiter (1982)[a] O: Donovan and Rossiter (1982)[b] O: Sweeney and Wyber (2002)[c]
Purchasing intentions/ spending	P: Sherman, Mathur, and Smith (1997) O: Spies, Hesse, and Loesch (1997)	N: Milliman (1982)[c] P: Sherman, Mathur, and Smith (1997)[c] O: Smith and Curnow (1966)[c]
Store visit duration	P: Donovan and Rossiter (1982) O: Sherman, Mathur, and Smith (1997) P: Spies, Hesse, and Loesch (1997)	P: Donovan and Rossiter (1982)[a] O: Donovan and Rossiter (1982)[b] P: Sherman, Mathur, and Smith (1997)[c] N: Smith and Curnow (1966)[c]
Affiliation (social interaction)	P: Donovan and Rossiter (1982) P: Dubé, Chebat, and Morin (1995) P: Sweeney and Wyber (2002)	P: Donovan and Rossiter (1982)[a] O: Donovan and Rossiter (1982)[b] P: Dubé, Chebat, and Morin (1995)[c] P: Sweeney and Wyber (2002)[c]
Unplanned spending	P: Donovan and Rossiter (1982) P: Donovan et al. (1994) P: Spies, Hesse, and Loesch (1997)	O: Donovan and Rossiter (1982)[a, b] O: Donovan et al. (1994)[a] N: Donovan et al. (1994)[b]
Unplanned extension of the store visit	P: Donovan et al. (1994)	O: Donovan et al. (1994)[a, b]
Satisfaction/ attitude	P: Sherman, Mathur, and Smith (1997) P: Spies, Hesse, and Loesch (1997) P: Wirtz and Bateson (1999) P: Yalch and Spangenberg (2000) P: Yoo, Park, and MacInnis (1998)	O: Sherman, Mathur, and Smith (1997)[c] O: Wirtz and Bateson (1999)[c] P: Yalch and Spangenberg (2000)[c]

[a] Pleasant environments only.
[b] Unpleasant environments only.
[c] The level of pleasantness in the environment is not controlled for.
Notes: Approach shopping behavior includes the other six variables, which are forms of approach. P = positive relationship, O = no relationship, and N = negative relationship.

unplanned spending, duration of the store visit, and social interaction. However, arousal effects are less consistent. For example, Sherman, Mathur, and Smith (1997) report that arousal increases purchasing intentions and spending, Milliman (1982) finds the opposite effect, and Smith and Curnow (1966) find no effect. Sherman, Mathur, and Smith find that arousal increases store visit duration, but Smith and Curnow find the opposite. Donovan and Rossiter (1982) replicate Sherman, Mathur, and Smith's results in pleasant retail environments only. Finally, in a study that involved different types of retail stores, Donovan and Rossiter find that arousal has a positive effect on most dependent variables in pleasant retail environments. However, Donovan and colleagues (1994) fail to confirm this relationship in a study conducted only in discount stores.

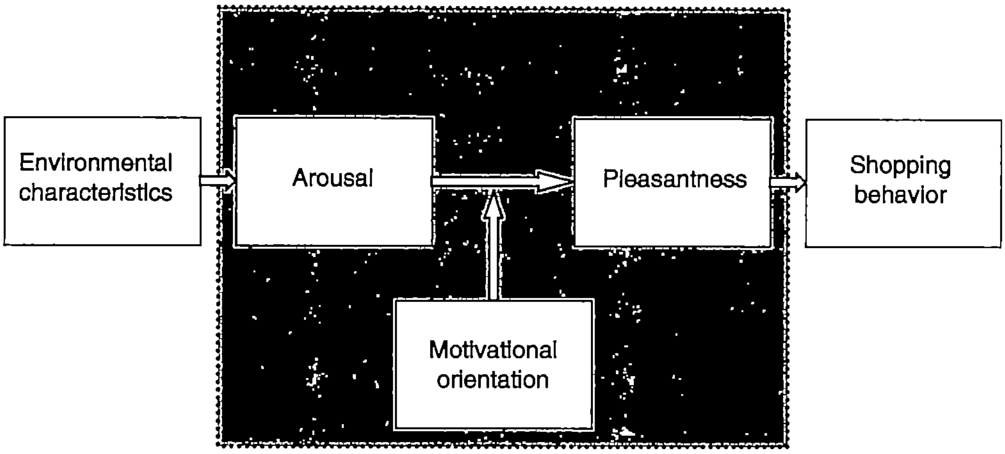
The overall pattern of results suggests that pleasantness has a consistently positive effect on shopping behavior, but the impact of arousal varies across studies. Thus, arousal effects may be moderated by a previously unidentified construct. In the next section, we propose a theoretical framework that resolves a number of inconsistencies in the research on arousal effects on shopping behavior, providing some managerially relevant insights into the appropriate design of store environments and retail Web sites.

Theoretical Framework

Our theoretical framework appears in Figure 1. The framework proposes that environmental characteristics affect consumer arousal, which in turn affects pleasantness and (through pleasantness) consumer shopping behaviors. Consumer motivational orientation moderates the relationship between arousal and pleasantness. When consumers have a recreational motivational orientation, arousal has a positive effect on pleasantness. Conversely, when consumers have a task-oriented motivational orientation, arousal has a negative effect on pleasantness.

The relationship between various environmental characteristics and arousal is well documented in the literature. For example, there is evidence that warm colors (Kueller and Mikellides 1993; Valdez and Mehrabian 1994), fast music tempo (Holbrook and Gardner 1993; Kellaris and

FIGURE 1
Theoretical Framework



Kent 1993), and complex environments (Berlyne 1960; Mehrabian and Russell 1974) increase arousal. The previously discussed studies in the marketing literature provide significant support for the proposed relationship between pleasantness and shopping behavior. Although we examine the relationship between pleasantness and shopping behavior in this research, the novel element in our framework (shown in the shaded area of Figure 1) is the proposed relationship among motivational orientation, arousal, and pleasantness. Thus, the theoretical discussion and the two experiments we describe subsequently focus on this element of the framework.

Constructs

In this section, we define the three constructs of primary interest: motivational orientation, arousal, and pleasantness. Then, in the next section, we develop a proposition about the relationship among these three constructs.

Motivational orientation. Various shopping motives were identified in the early literature (Stone 1954; Tauber 1972; for a review, see Westbrook and Black 1985). However, subsequent research found that two fundamental motivational orientations underlie the different shopping motives. A summary of the empirical studies over the past 20 years that investigate the two underlying shopping motivational orientations appears in Table 2.

The first motivational orientation (e.g., economic, utilitarian) involves consumers engaging in shopping out of necessity to obtain needed products, services, or information with little or no inherent satisfaction derived from the shopping activity itself. We refer to this motivational orientation as the "task-oriented motivational orientation." The other motivational orientation (e.g., recreational, hedonic) describes consumers engaging in shopping to derive inherent satisfaction from the shopping activity itself. In this case, the shopping activity is freely chosen, and there is no need to engage in it. We refer to this motivational orientation as the "recreational motivational orientation." The two fundamental motivational orientations identified in the shopping-behavior literature are consistent with the fundamental motivational orientations identified in psychology (Apter 1982; Deci and Ryan 1985).

Arousal. The arousal construct is used to designate two somewhat distinct phenomena in psychology. "Objective arousal" is defined as the release of energy collected in the tissues (Duffy 1962), a variable representing the level of nonspecific physiological response (Cacioppo, Berntson, and Crites 1996). "Subjective arousal" is defined as the subjective experience of energy mobilization (Russell and Feldman Barrett 1999). Objective and subjective arousal are moderately correlated (Thayer 1970). Mehrabian and Russell's (1974) framework involves subjective arousal. Mehrabian and Russell define arousal as an affective dimension ranging from sleep to frantic excitement and use a self-report scale to measure it. Our framework also involves subjective arousal because it is the subjective state that consumers can experience as pleasant or unpleasant.

Pleasantness. "Pleasantness" is defined as the hedonic valence (pleasant or unpleasant) of the affective response to a stimulus; it is based on the extent to which the stimulus (the object of the affective response) enables people to achieve their salient goals. Stimuli that facilitate goal achievement are experienced as pleasant, whereas stimuli that impede goal achievement are experienced as unpleasant (for a review, see Clore, Schwarz, and Conway 1994).

Motivational Orientation as Moderator

Because task-oriented consumers derive satisfaction from the outcome of the shopping activity (the acquisition of the needed product, service, or information) rather than from the activity itself, their focus is on efficiently completing the shopping activity and obtaining its outcome with minimum expense of energy. Such consumers would find that the high energy demands in high-arousal environments require more effort to complete the shopping activity and therefore would find such environments to be unpleasant. In contrast, the recreational motivational orientation involves pursuing shopping to derive intrinsic rewards from the shopping activity itself. Such consumers desire rich experi-

TABLE 2
Task-Oriented and Recreational Shopping Motivational Orientations: Conceptualizations in the Shopping-Behavior Literature

Article	Motivational Orientation Labels	Definitions	
		Task Oriented	Recreational
Bellenger and Korgaonkar (1980)	Economic shoppers Recreational shoppers	Economic shoppers: Consumers who have a high opportunity cost for shopping activities. Economic shoppers feel that shopping is not an enjoyable use of their time (p. 79).	Recreational shoppers: Consumers who have a low opportunity cost for shopping activities Recreational shoppers feel that shopping is an enjoyable use of their time (p. 79).
Westbrook and Black (1985)	Economic shoppers Social shoppers	Economic shoppers: Consumers who engage in shopping primarily to acquire (a) needed product(s) (p. 85).	Social shoppers: Consumers who engage in shopping primarily to satisfy needs unrelated to the acquisition of needed products (p. 85).
Dawson, Bloch, and Ridgway (1990)	Product-oriented motives Experiential motives	Product-oriented shopping motives: Consumers visit stores to make needed purchases or acquire needed product information (p. 409).	Experiential shopping motives: Consumers visit stores to derive pleasure from the visit itself (p. 410).
Babin, Darden, and Griffin (1994)	Utilitarian shopping value Hedonic shopping value	Utilitarian shopping value: Derived from task completion, the achievement of an intended outcome that the consumer is pursuing out of necessity (pp. 645–46).	Hedonic shopping value: Derived from the spontaneous hedonic responses elicited in the course of shopping activities (pp. 645–46).
Reynolds and Beatty (1999)	Shopping reluctance/apathy Shopping enjoyment	Shopping reluctance/apathy: Satisfaction is not derived from the shopping activity itself (p. 511).	Shopping enjoyment: Satisfaction is derived from the shopping activity itself (p. 511).

ences from shopping and therefore would find the high energy demands in high-arousal environments to be pleasant; conversely, they would view environments that call for less energy mobilization as unpleasant.

The following scenario illustrates this phenomenon: A businesswoman is likely to have a task-oriented motivational orientation when she feels that, while on a business trip to New York, she must find a gift to take to her assistant's son. Because the gift-buying activity is not inherently satisfying, she will try to complete the task as efficiently as possible. She will find the high-arousal environment (e.g., loud music, bright lights, action displays) in the Times Square Toys "R" Us store to be unpleasant because it will require her to mobilize more energy to complete her shopping task. In contrast, if the businesswoman were interested in some recreational shopping before departing from New York City, she would seek to derive some intrinsic satisfaction from the shopping activity itself. As a result, she would find the high-arousal environment at the Times Square Toys "R" Us to be pleasant because it would enhance her shopping experience. On the basis of the preceding discussion, we propose the following:

> P_1 Motivational orientation moderates the effect of arousal on pleasantness: Arousal (a) decreases pleasantness for consumers with a task-oriented motivational orientation and (b) increases pleasantness for consumers with a recreational motivational orientation.

The proposed moderating effect of motivational orientation resolves some of the inconsistent findings in the shopping-behavior literature we reviewed previously. For example, Milliman (1982) finds that arousal decreases purchasing intentions and spending in supermarkets. In contrast, Sherman, Mathur, and Smith (1997) find that arousal has a positive effect on purchasing intentions and spending in mall-based fashion stores. Few consumers find grocery shopping intrinsically satisfying (Levy and Weitz 2004). Thus, we can assume that consumers shopping for groceries are likely to have a task-oriented motivational orientation, which explains Milliman's results. Conversely, consumers often visit mall-based fashion stores as a form of recreation. The assumption that the consumers in Sherman, Mathur, and Smith's study were predominantly recreationally oriented would explain their results.

Empirical Research

Procedure

We used the same procedure in both Study 1 and Study 2. The participants were undergraduate students from a state university in the southeastern United States who volunteered to participate in the studies in return for extra credit in introductory marketing classes. Participants were first induced to adopt either a recreational motivational orientation or a task-oriented motivational orientation. Then, participants visited a high-arousal or a low-arousal shopping environment that was similar to shopping over the Internet. Next, participants indicated the pleasantness of their experience in the environment and their shopping-behavior intentions in relation to the environment. The studies concluded with manipulation checks.

Measures

All studies used the same measures to assess the dependent variables (pleasantness and shopping-behavior intentions) and the effectiveness of the arousal and motivation manipulations. (The complete scales, with descriptive statistics, appear in the Appendix.) We assessed pleasantness with four items based on Mehrabian and Russell's (1974) pleasantness scale. We could not use the entire scale, because three of the anchor adjectives (hopeful, despairing, and melancholic) signify intense emotions that are not typically felt in retail contexts, and one of the items (relaxed versus bored) implies recreational motivation. Mehrabian and Russell's arousal scale was used for the arousal manipulation check. We developed a four-item scale for the motivation manipulation check. A six-item scale based on the work of Donovan and Rossiter (1982) was used to assess participants' shopping-behavior intentions.

We assessed the psychometric properties of the measures with confirmatory factor analysis. To ensure an adequate sample size for the confirmatory factor analysis, all observations from Studies 1 and 2 were combined into a single sample, and the confirmatory factor analysis was conducted on the combined sample. The error variances of the three negatively worded behavioral intention items were allowed to correlate because of shared method. Although the chi-square test was statistically significant ($\chi^2(164) = 495.28$, $p < .01$), the comparative fit index (.95), the Tucker–Lewis index (.94), and the root mean square error of approximation (.08) indicated an acceptable fit (Bagozzi and Baumgartner 1994; Bagozzi and Yi 1988; Browne and Cudeck 1993). The scale reliabilities had acceptable psychometric properties: The coefficient alpha values were between .81 and .90, and the composite reliabilities were between .50 and .66 (Bagozzi and Baumgartner 1994).

We used two different methods to assess discriminant validity (Anderson and Gerbing 1988; Fornell and Larcker 1981). According to Fornell and Larcker (1981), discriminant validity between two factors is established when the squared correlation between them is less than the mean item communality for each one. All six squared interfactor correlations were less than the mean item communality of either factor, demonstrating that the four measures have discriminant validity. Anderson and Gerbing's (1988) method involves chi-square difference tests. Two models are estimated for each pair of two factors. One of the models is a one-factor model; all items are set to load on a single factor. In the other model, each item is set to load on its respective factor. A chi-square difference test is then performed to determine whether the two-factor model explains significantly more variance than the one-factor model. Discriminant validity between the two factors is established if the two-factor model explains significantly more variance than the one-factor model. We performed six chi-square difference tests. All six tests resulted in significant chi-square differences, in support of the discriminant validity of the measures.

Study 1

We conducted Study 1 as a 2 (motivational orientation: task oriented versus recreational) × 2 (arousal: high versus low) × 2 (replicates) between-subjects experiment. The 166 participants in the study were asked to view a picture of a retail store ambience on individual computer screens and to imagine visiting the store with a task-oriented or recreational motivation.

Manipulation of Motivational Orientation

Participants were randomly assigned to read one of the following two scenarios about a hypothetical shopping experience:

> You are going on a camping trip this weekend, and you realize that you don't have enough suitable T-shirts for the trip. As a result, you decide to purchase at least one more T-shirt. You drive to a store that sells T-shirts. All you want to do in the store is find one or more suitable T-shirts for your trip and leave. (*Task-oriented motivation treatment*)
>
> It's just past noon on a Saturday, and none of your friends are around. It's pouring rain, so you can't do anything outdoors. You find what's on TV too dull to watch. You feel very, very bored. You decide to visit some stores to relieve the sense of boredom. (*Recreational motivation treatment*)

To increase the salience of the respective motivational orientation, participants were asked to write a story five or more sentences long about themselves in a similar situation.

Manipulation of Arousal

The arousal manipulation involved varying three visual elements of the stimulus shopping environment: (1) complexity (the number of nonredundant elements in an environment; Berlyne 1971), (2) color warmth (the warmth–coolness color continuum has been identified as red, yellow, green, and blue; Erwin et al. 1961; Jacobs and Hustmyer 1974), and (3) color saturation (the subjective experience of a wavelength's spectral purity; Hogg 1969). Higher complexity (Berlyne 1960; Mehrabian and Russell 1974), warmer colors (Kueller and Mikellides 1993; Valdez and Mehrabian 1994), and higher color saturation (Valdez and Mehrabian 1994) all increase arousal. There were two replicates of the arousal manipulation: two high-arousal environments and two low-arousal environments. Participants were randomly assigned to view one of the four environments. All participants were exposed to the treatment environment for the same amount of time (i.e., 20 seconds).

Manipulation Checks

We conducted a 2 (motivational orientation) × 2 (arousal) × 2 (replicates) analysis of variance (ANOVA) to test the effectiveness of the two manipulations. Motivational orientation, arousal, and the replicates, as well as all interactions among these three variables, were the independent variables. The motivation and arousal manipulation check measures were the dependent variables. The participants in the recreational motivation condition rated their motivational orientation as significantly more recreational than the participants in the task-oriented motivation condition ($F(1, 158) = 70.56$, $p < .01$; $M_{RC} = 5.92$ versus $M_{TO} = 3.82$). There were no other significant main or interaction effects on the motivation manipulation check ($F(1, 158) \leq 2.83$, $p > .09$). The arousal manipulation had a significant main effect on the arousal manipulation check ($F(1, 158) = 30.86$, $p < .01$; $M_H = 5.58$ versus $M_L = 4.59$). There were no other significant effects on the arousal manipulation check ($F(1, 158) \leq 2.02$, $p > .15$).

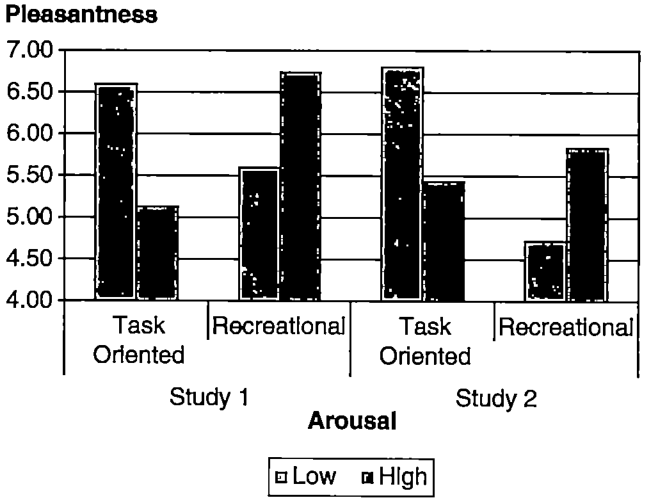
Results

Pleasantness. We conducted a 2 (motivational orientation) × 2 (arousal) × 2 (replicates) ANOVA to test the hypothesized interaction effect between arousal and motivation on pleasantness. We found that motivational orientation moderates the effect of arousal on pleasantness, producing a significant interaction between arousal and motivation ($F(1, 158) = 28.08$, $p < .01$). Neither motivational orientation ($F(1, 158) = 1.64$, $p > .20$; $M_{RC} = 6.17$ versus $M_{TO} = 5.86$) nor arousal ($F(1, 158) = .40$, $p > .20$; $M_L = 6.09$ versus $M_H = 5.94$) had a significant main effect on pleasantness. There were no other significant effects ($F(1, 158) \leq 1.07$, $p > .20$).

We conducted a 2 (arousal) × 2 (replicates) ANOVA separately in each motivation condition to test whether the significant interaction between arousal and motivation was a significant crossover effect. The participants in the task-oriented motivation condition found the low-arousal environment to be significantly more pleasant than the high-arousal environment ($F(1, 80) = 21.47$, $p < .01$; $M_L = 6.59$ versus $M_H = 5.13$). The participants in the recreational motivation condition found the high-arousal environment to be significantly more pleasant than the low-arousal environment ($F(1, 78) = 9.12$, $p < .01$; $M_H = 6.75$ versus $M_L = 5.60$) (see Figure 2). There were no other significant effects in these analyses ($F \leq 1.22$, $p > .20$).

Shopping behavior. We estimated the same ANOVA models to test whether arousal and motivation have an interactive effect on shopping-behavior intentions. We found a significant interaction between arousal and motivation ($F(1, 158) = 14.80$, $p < .01$). Neither motivational orientation ($F(1, 158) = .18$, $p > .20$; $M_{RC} = 5.94$ versus $M_{TO} = 5.84$) nor arousal ($F(1, 158) = .28$, $p > .20$; $M_L = 5.95$ versus $M_H = 5.83$) had a significant main effect on behavioral intentions. Arousal had a significant, negative effect on behavioral intentions in the task-oriented motivation condition ($F(1, 80) = 16.25$, $p < .01$; $M_L = 6.36$ versus $M_H = 5.32$) and a significant, positive effect on behavioral intentions in the recreational motivation condition ($F(1, 78) = 3.81$, $p = .05$; $M_H = 6.34$ versus $M_L = 5.55$). There were no other significant effects in these analyses ($F \leq 1.34$, $p > .20$).

FIGURE 2
Pleasantness as a Function of Motivational Orientation and Arousal



To test whether pleasantness mediates the interactive effect between arousal and motivation on shopping-behavior intentions, we conducted the mediation analysis recommended by Baron and Kenny (1986). We coded the two levels of each independent variable (motivational orientation, arousal, and the replicates) as 1 or –1. The behavioral intentions variable was the dependent variable. We estimated two regression models:

$$
\begin{aligned}
(1)\quad \text{Behavioral intentions} = {} & \beta_0 + \beta_1 \times \text{motivational orientation} \\
& + \beta_2 \times \text{arousal} + \beta_3 \times \text{replicates} \\
& + \beta_4 \times \text{motivational orientation} \\
& \times \text{arousal} + \beta_5 \\
& \times \text{motivational orientation} \\
& \times \text{replicates} + \beta_6 \times \text{arousal} \\
& \times \text{replicates} + \beta_7 \\
& \times \text{motivational orientation} \times \text{arousal} \\
& \times \text{replicates}.
\end{aligned}
$$

$$
\begin{aligned}
(2)\quad \text{Behavioral intentions} = {} & \beta_0 + \beta_1 \times \text{motivational orientation} \\
& + \beta_2 \times \text{arousal} + \beta_3 \times \text{replicates} \\
& + \beta_4 \times \text{motivational orientation} \\
& \times \text{arousal} + \beta_5 \\
& \times \text{motivational orientation}
\end{aligned}
$$

$\times$ replicates + $\beta_6 \times$ arousal

$\times$ replicates + β_7

$\times$ motivational orientation $\times$ arousal

$\times$ replicates + $\beta_8 \times$ pleasantness.

Evidence in support of mediation arises if β_4 is statistically significant in Model 1 but not statistically significant in Model 2. The results support the hypothesized mediating effect of pleasantness. The interaction between arousal and motivation was significant in Model 1 (β_4 = .29, t(158) = 3.85, $p < .01$) but was not significant in Model 2 (β_4 = .05, t(157) = .74, $p > .40$). Pleasantness had a significant, positive effect on behavioral intentions in Model 2 (β_8 = .62, t(157) = 9.51, $p < .01$).

Study 2

The objectives of Study 2 were (1) to replicate the results we obtained in Study 1 with different and more realistic manipulations of arousal and motivational orientation and (2) to provide further insight into the processes underlying the relationship between arousal and pleasantness by collecting free-response data from participants. The study was a 2 (motivational orientation: task oriented versus recreational) × 2 (arousal: high versus low) between-subjects experiment.

Subjects and Procedure

The 161 participants were told that a music retailer was planning to open a music store in town. Participants were asked to look at and evaluate a store concept as prospective customers. On individual computer screens, participants were shown an interior view of a music store that sells CDs. Before the music store was displayed on the screen, onscreen instructions informed participants that more detailed information could be obtained about the CDs shown in the store by positioning the mouse over the merchandise. A mouse click on a CD revealed a CD information window that displayed the composer's portrait, a brief description of the CD, and the CD's price. Participants could return to the store view window by pressing a button in the CD information window. While shopping in the store, participants heard music through earphones. So that participants would become familiar with the setting, they were asked to complete a trial task—namely, to open a product's information window in a trial store and find out the product's price. The trial store was different from the treatment store and was the same for all participants.

Manipulation of motivational orientation. In the task-oriented motivation condition, participants were asked to imagine themselves in the situation described in the following scenario:

> You are planning to meet some friends at a movie theater. Your mother asks you to stop at a certain nearby music store (before going to the movie theater) and check the prices of two classical music CDs. Your mother intends to purchase one of those two CDs as a gift for an elderly couple, but she hasn't made up her mind which CD to choose. Your mother gives you the composers' names and album titles: Beethoven's *9 Symphonien* and Chopin's *The Piano Works*. Your task is to go to the store, find the CDs, and remember their prices.

Then participants were specifically instructed to locate the two CDs indicated in the scenario and uncover their prices. There was no time limit on the visit to the store; participants could leave the store at any time by pressing an "exit" button in the store view window. Participants were also informed that at the end of the session, they would be asked to identify the two CD prices in a multiple-choice question. This element in the instructions was necessary so that participants would view identifying the CD prices as a task rather than as a recreational activity that may or may not be completed. However, participants were informed that there would be no penalty or other adverse consequences for those who failed to identify the two prices correctly.

The participants in the recreational motivation condition were asked to imagine themselves in the situation described in the following scenario:

> It is spring break. You have to stay in town while all your friends have left town. It is a weekend. You are at home, and you feel very, very bored. It's pouring rain, so you can't do anything outdoors. You find what's on TV too dull to watch. So you decide to visit some stores to relieve the sense of boredom.

Participants were specifically instructed to visit the music store "to relieve boredom (as when you feel very, very bored and look for diversion)." They were also informed that more detailed information could be obtained about the CDs shown in the store and that they could (but did not have to) access that information. Participants were also informed that there was no time limit on their visit to the store; they could leave the store at any time by pressing an "exit" button in the store view window. Notepaper was available to each participant in both motivation conditions.

Manipulation of arousal. We used music tempo and volume to manipulate arousal. The high-arousal treatment involved faster and louder music (the first 60 seconds of Tchaikovsky's "Capriccio Italien in A"), and the low-arousal treatment involved slower and softer music (the first 60 seconds of Camille Saint-Saens' "The Carnival of the Animals: The Swan"). Participants who remained in the store for more than a minute heard the same musical selection from the beginning.

Manipulation checks. We conducted a 2 (motivational orientation) × 2 (arousal) ANOVA to test the effectiveness of the two manipulations. Motivational orientation, arousal, and their interaction were the independent variables. The motivation and arousal manipulation check measures were the dependent variables. The participants in the recreational motivation condition rated their motivational orientation as significantly more recreational than the participants in the task-oriented motivation condition (F(1, 157) = 123.47, $p < .01$; $M_{RC} = 5.84$ versus $M_{TO} = 3.58$). There were no other significant main or interaction effects on the motivation manipulation check (F(1, 157) = .77, $p > .20$). The arousal manipulation had a main effect on the arousal manipulation check (F(1, 157) = 129.39, $p < .01$; $M_H = 6.02$ versus $M_L = 3.80$). There were no other significant main or interaction effects on the arousal manipulation check ($F(1, 157) = 1.55$, $p > .20$).

Results

Pleasantness. We conducted the same 2 (motivational orientation) × 2 (arousal) ANOVA to test the hypothesized moderating effect of motivational orientation on the relationship between arousal and pleasantness. We found that motivational orientation significantly moderated the effect of arousal on pleasantness ($F(1, 157) = 20.25$, $p < .01$). Motivational orientation also had a main effect on pleasantness: The participants in the task-oriented motivation condition indicated more pleasant affect than the participants in the recreational motivation condition ($F(1, 157) = 9.44$, $p < .01$; $M_{TO} = 6.13$ versus $M_{RC} = 5.28$). Arousal had no significant main effect on pleasantness ($F(1, 157) = .20$, $p > .20$; $M_L = 5.77$ versus $M_H = 5.64$).

We examined the effect of arousal on pleasantness separately in each motivation condition to test for a significant crossover interaction. The participants in the task-oriented motivation condition found the low-arousal environment to be significantly more pleasant than the high-arousal environment ($F(1, 79) = 13.53$, $p < .01$; $M_L = 6.81$ versus $M_H = 5.44$). The participants in the recreational motivation condition found the high-arousal environment to be significantly more pleasant than the low-arousal environment ($F(1, 78) = 7.47$, $p < .01$; $M_H = 5.84$ versus $M_L = 4.73$) (see Figure 2).

Free-response process measures. Participants were asked a free-response question about the store environment to examine their thought processes in the different conditions and to test our theoretical rationale. After experiencing the treatment store environment (the arousal manipulation), in their own words, participants answered the question, "Why did (didn't) you like the store environment?" Two coders independently coded participants' responses into categories. Each coder independently coded a subset of responses to establish a set of relevant categories. Then, the two coders met and, through discussion, determined the set of categories to be used for classifying participants' responses. Next, the coders used the developed categorization scheme and independently coded all participant responses. Coder agreement was calculated as the percentage of responses that both coders independently classified into the same category. The agreement rate was 92%. Disagreements were resolved through discussion. On average, participants listed 1.81 reasons for their store experience. We used one reason per participant in the analysis—the reason indicated as the most important. If a participant listed more than one reason and did not explicitly indicate which reason was the most important one, we used the reason listed first. Table 3 shows the reasons listed in each arousal × motivation condition and the frequency with which each reason occurs in the condition.

The free-response data are consistent with our explanation for the interactive effect between arousal and motivation. Task-oriented participants liked the low-arousal store environment because it allowed them to complete their shopping efficiently, but they did not like the high-arousal environment because it created inefficiencies. Recreational

TABLE 3
Reasons for Participants' Store Environment Experience

Experience Reason	Frequency	
	n	%
Task-Oriented Low-Arousal Condition (n = 39) (found to be pleasant in general)		
1. Efficiency (effortlessness, expediency)	31	79.5
2. Boredom (limited opportunity to engage in activities)	3	7.7
3 Demotivation (no motivation to engage in any activity)	2	5.1
Task-Oriented High-Arousal Condition (n = 41) (found to be unpleasant in general)		
1, Inefficiency (effortfulness, extra time needed)	26	63.4
2 Individual difference (dispositional) reasons	4	9.8
3. Rich experience (ample opportunity to engage in activities)	3	7.3
4. Pleasantness (nonspecific reasons, such as "because it's pleasant")	3	7.3
5. Efficiency (effortlessness, expediency)	3	7.3
Recreational Low-Arousal Condition (n = 40) (found to be unpleasant in general)		
1. Boredom (limited opportunity to engage in activities)	30	75.0
2. Efficiency (effortlessness, expediency)	3	7.5
3. Individual difference (dispositional) reasons	2	5.0
4. Pleasantness (nonspecific reasons, such as "because it's pleasant")	2	5.0
Recreational High-Arousal Condition (n = 40) (found to be pleasant in general)		
1. Rich experience (ample opportunity to engage in activities)	19	47.5
2. Pleasantness (nonspecific reasons, such as "because it's pleasant")	7	17.5
3. Individual difference (dispositional) reasons	6	15.0
4. Unpleasantness (nonspecific reasons, such as "because it's unpleasant")	4	10.0
5. Inefficiency (effortfulness, extra time needed)	3	7.5

Notes: Reasons listed only once in a condition are not shown.

participants liked the high-arousal store environment because it created rich experiences, but they disliked the low-arousal environment because it involved limited opportunities to engage in activities.

Shopping behavior. A 2 (motivational orientation) × 2 (arousal) ANOVA was estimated to test whether arousal and motivational orientation affect shopping-behavior intentions. A significant interaction between arousal and motivation was obtained ($F(1, 157) = 6.10$, $p < .05$). Neither motivational orientation ($F(1, 157) = 1.03$, $p > .20$; $M_{TO} = 5.42$ versus $M_{RC} = 5.13$) nor arousal ($F(1, 157) = 1.17$, $p > .20$; $M_H = 5.43$ versus $M_L = 5.11$) had a significant main effect on behavioral intentions. We also examined the effect of arousal on behavioral intentions separately in each motivation condition. Arousal had a significant, positive effect on behavioral intentions in the recreational motivation condition ($F(1, 78) = 6.20$, $p < .05$; $M_H = 5.65$ versus $M_L = 4.61$) but did not have a significant effect in the task-oriented motivation condition ($F(1, 79) = .98$, $p > .20$; $M_L = 5.63$ versus $M_H = 5.22$).

To test for the hypothesized mediating effect of pleasantness, we conducted the same type of mediation analysis as that in Study 1. The results were consistent with expectations. The interaction between arousal and motivation was significant in the model that did not include pleasantness ($\beta_4 = .33$, $t(157) = 2.47$, $p < .05$), but it was not significant in the model that included pleasantness ($\beta_4 = -.09$, $t(156) = -.83$, $p > .40$). Pleasantness had a significant, positive effect on behavioral intentions in Model 2 ($\beta_8 = .73$, $t(156) = 11.94$, $p < .01$). Because arousal had no effect on behavioral intentions in the task-oriented motivation condition, we also conducted a mediation analysis only for the recreational motivation condition. Arousal had a significant, positive effect on behavioral intentions in the model that did not include pleasantness ($\beta_4 = .27$, $t(78) = 2.49$, $p < .05$), but it had no effect in the model that included pleasantness ($\beta_4 = .07$, $t(77) = .85$, $p > .30$). Pleasantness had a significant, positive effect on behavioral intentions in the latter model ($\beta_8 = .67$, $t(77) = 7.87$, $p < .01$).

General Discussion

The two studies provide consistent support for the proposition that motivational orientation moderates the effect of arousal on pleasantness. In both studies, there is a significant crossover interaction, in which both simple effects are significant. The generalizability of the results is strengthened by the use of different manipulations of arousal and motivational orientation in the two studies. As we proposed, task-oriented consumers view shopping primarily as a means for obtaining a needed outcome (product, service, or information) and wish to complete their shopping as efficiently as possible; therefore, task-oriented consumers find high-arousal retail environments to be unpleasant. Conversely, recreational consumers derive inherent satisfaction from the shopping activity itself and therefore like high-arousal retail environments that create rich shopping experiences. The free-response data we collected in Study 2 supports our explanation for the interactive effect between arousal and motivational orientation.

In addition, the two studies provide evidence that is consistent with the results pertaining to shopping behavior obtained in the marketing literature (Table 1). Arousal and motivational orientation have an interactive effect on shopping behavior, and this interactive effect is mediated by pleasantness. Note, however, that the interactive effect between arousal and motivation was not a crossover effect in Study 2; arousal had no significant effect on shopping-behavior intentions in the task-oriented condition in that study.

Limitations

The two studies involve participants engaging in a simulated shopping experience. Although Study 2 creates an experience similar to the experience that consumers would have when shopping on the Internet, the studies do not capture many elements of the experience that consumers have when shopping in stores. Thus, there is a need to provide more information about the generalizability of the results by carrying out either controlled experiments or correlational studies in stores.

Managerial Implications

This research indicates that the level of excitement retailers should create in their stores depends on the shopping motivation of their customers. Retailers can infer the motivational orientation of their current and prospective customers from the retail offering (e.g., school supplies versus books, dental care versus cruises), the day of the week and the time of day (e.g., a Monday morning versus a Saturday afternoon), the time of the year (a nonholiday versus a holiday season), and the retailer's location (e.g., a business district versus a theme park). After a retailer determines the dominant motivational orientation of its customers, our framework suggests that the retailer should create a high-arousal environment for recreational consumers and a low-arousal environment for task-oriented consumers.

Retailers can use different elements of the store environment to influence customer arousal levels (e.g., music, color, complexity of the store layout and the merchandise presentation). For example, if a grocery store chain finds out that its customers are predominantly task oriented, it should use store designs with less-saturated cooler colors (e.g., light blues) and even associate such colors with its brand name. It should not use intense orange or red colors. The stores should have a simple layout and merchandise presentation. In contrast, the customers of a sporting-goods retailer may be predominantly recreationally oriented. This retailer should design its stores with a more complex layout and merchandise presentation and regularly reorganize the store and the merchandize presentation. The retailer should also use warm and highly saturated colors throughout its stores and in its brand name.

When customers' dominant motivational orientation varies in time, the retailer faces a more challenging task in designing the store environment. The elements of the store environment that are more difficult to modify frequently (e.g., the store's layout and wall color) should be designed to induce moderate arousal (i.e., intermediate levels of complexity, color warmth, and color saturation). More variable elements, such as the background music, can be used to increase arousal at the times when customers are predominantly recreationally oriented and to decrease arousal when customers are predominantly task oriented.

Sometimes customers' dominant motivational orientation varies across the different store departments. In such cases, retailers are advised to design each department in such a way as to complement the dominant motivational orientation of the customers shopping there. For example, a consumer electronics retailer may create a low-arousal environment in the accessories area to accommodate customers (who typically have a task-oriented motivational orientation when shopping for accessories) but create a high-arousal environment in the home-entertainment centers, which are typically visited by recreationally oriented consumers.

Although our research and review of the literature focus on environmental effects in retail stores, the framework can be applied to consumer experiences with other marketing stimuli (e.g., packages and labels, advertising messages, Web sites). For example, according to our framework, Amazon.com should serve more complex, high-arousal Web sites to customers who indicate that they have come to the Web site to browse, but it should offer simpler, low-arousal Web sites to customers trying to find a specific book.

Conclusion

Our theoretical framework identifies and examines a situational variable, motivational orientation, that moderates the relationship between arousal and pleasantness. By identifying the antecedents that determine the pleasantness of people's experiences while engaging in shopping and consumption, the framework offers an explanation for important aspects of consumer behavior and furnishes guidelines for effectively managing retail and consumption environments.

Appendix
Measurement Scales

Dependent Variables

Pleasantness (scored on a nine-point semantic differential scale)

(Study 1) In this store, I would feel:

(Study 2) Throughout my visit to the concept music store, I felt:

- Displeased versus pleased; satisfied versus dissatisfied (reversed); pleasant versus unpleasant (reversed); unhappy versus happy.

(Study 1: M = 6.05, SD = 1.70, α = .856; Study 2: M = 5.70, SD = 1.88, α = .912; Studies 1 and 2: M = 5.88, SD = 1.80, α = .887)

Behavioral Intentions (scored on a nine-point Likert scale, anchored by "strongly disagree" and "strongly agree")

What would you generally do if this store opens in this town or your hometown?

- I would enjoy shopping in this store.
- I would avoid ever having to return to this store. (reversed)
- I would want to avoid looking around or exploring this environment. (reversed)
- This is a place where I might try to avoid other people, and avoid having to talk to them. (reversed)
- I would be willing to buy things at this store.
- I would be willing to recommend this store to my friends.

(Study 1: M = 5.94, SD = 1.60, α = .884; Study 2: M = 5.27, SD = 1.89, α = .907; Studies 1 and 2: M = 5.61, SD = 1.77, α = .900)

Manipulation Checks

Arousal (scored on a nine-point semantic differential scale)

(Study 1) In this store, I would feel:

(Study 2) Throughout my visit to the concept music store, I felt:

- Relaxed versus stimulated; excited versus calm (reversed); frenzied versus sluggish (reversed); dull versus jittery; wide awake versus sleepy (reversed); unaroused versus aroused.

(Study 1: M = 5.11, SD = 1.23, α = .784; Study 2: M = 4.92, SD = 1.66, α = .886; Studies 1 and 2: M = 5.02, SD = 1.46, α = .848)

Motivational Orientation (scored on a nine-point Likert scale, anchored by "strongly disagree" and "strongly agree")

(Study 1) On this shopping occasion, I would primarily want:

(Study 2) In the store, I primarily wanted:

- To have fun; to get things done (reversed); to be task focused (reversed); to relieve boredom.

(Study 1: M = 4.87, SD = 1.91, α = .850; Study 2: M = 4.70, SD = 1.72, α = .771; Studies 1 and 2: M = 4.79, SD = 1.82, α = .808)

REFERENCES

Anderson, James C. and David W. Gerbing (1988), "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin*, 103 (May), 411–23.

Apter, Michael J. (1982), *The Experience of Motivation: The Theory of Psychological Reversals.* New York: Academic Press.

Babin, Barry J., William R. Darden, and Mitch Griffin (1994), "Work and/or Fun: Measuring Hedonic and Utilitarian Shopping Value," *Journal of Consumer Research*, 20 (March), 644–56.

Bagozzi, Richard P. and Hans Baumgartner (1994), "The Evaluation of Structural Equation Models and Hypothesis Testing," in *Principles of Marketing Research*, Richard P. Bagozzi, ed. Cambridge, MA: Blackwell Publishers, 386–422.

——— and Youjae Yi (1988), "On the Evaluation of Structural Equation Models," *Journal of the Academy of Marketing Science*, 16 (Spring), 74–94.

Baker, Julie, Michael Levy, and Dhruv Grewal (1992), "An Experimental Approach to Making Retail Store Environmental Decisions," *Journal of Retailing*, 68 (Winter), 445–60.

Baron, Reuben M. and David A. Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (December), 1173–82.

Bellenger, Danny N. and Pradeep K. Korgaonkar (1980), "Profiling the Recreational Shopper," *Journal of Retailing*, 56 (Fall), 77–92.

Berlyne, Daniel E. (1960), *Conflict, Arousal, and Curiosity*. New York: McGraw-Hill.

——— (1971), *Aesthetics and Psychobiology*. New York: Appleton-Century-Crofts.

Browne, Michael W. and Robert A. Cudeck (1993), "Alternative Ways of Assessing Model Fit," in *Testing Structural Equation Models*, Kenneth A. Bollen and J. Scott Long, eds. Newberry Park, CA: Sage Publications, 136–62.

Cacioppo, John T., Gary G. Berntson, and Stephen L. Crites (1996), "Social Neuroscience: Principles of Psychophysiological Arousal and Response," in *Social Psychology: Handbook of Basic Principles*, E. Tory Higgins and Arie W. Kruglanski, eds. New York: Guilford Press, 72–101.

Clore, Gerald L., Norbert Schwarz, and Michael Conway (1994), "Affective Causes and Consequences of Social Information Processing," in *Handbook of Social Cognition*, Vol. 1, Robert S. Wyer and Thomas K. Srull, eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 323–417.

Da Cunha, Marcus, Amitav Chakravarti, and Barton Weitz (2003), "Consumer Search, Design Preferences and Internet Search Behaviors," working paper, Marketing Department, University of Florida.

Dawson, Scott, Peter H. Bloch, and Nancy M. Ridgway (1990), "Shopping Motives, Emotional States, and Retail Outcomes," *Journal of Retailing*, 66 (Winter), 408–427.

Deci, Edward L. and Richard M. Ryan (1985), *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum Press.

Donovan, Robert J. and John R. Rossiter (1982), "Store Atmosphere: An Environmental Psychology Approach," *Journal of Retailing*, 58 (Spring), 34–57.

———, ———, Gilian Marcoolyn, and Andrew Nesdale (1994), "Store Atmosphere and Purchasing Behavior," *Journal of Retailing*, 70 (Fall), 283–94.

Dubé, Laurette, Jean-Charles Chebat, and Sylvie Morin (1995), "The Effects of Background Music on Consumers' Desire to Affiliate in Buyer-Seller Interactions," *Psychology & Marketing*, 12 (July), 305–319.

Duffy, Elizabeth (1962), *Activation and Behavior*. New York: John Wiley & Sons.

Erwin, C. William, Marvin Lerner, Norman J. Wilson, and William P. Wilson (1961), "Some Further Observations on the Photically Elicited Arousal Response," *Electroencephalography and Clinical Neurophysiology*, 13 (June), 391–94.

Fornell, Claes and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39–50.

Hogg, James (1969), "A Principal Components Analysis of Semantic Differential Judgments of Single Colors and Color Pairs," *Journal of General Psychology*, 80 (January), 129–40.

Holbrook, Morris B. and Meryl P. Gardner (1993), "An Approach to Investigating the Emotional Determinants of Consumption Durations: Why Do People Consume What They Consume for as Long as They Consume It?" *Journal of Consumer Psychology*, 2 (April), 123–42.

Hui, Michael K. and John E.G. Bateson (1991), "Perceived Control and the Effects of Crowding and Consumer Choice on the Service Experience," *Journal of Consumer Research*, 18 (September), 174–84.

Jacobs, Keith W. and Frank E. Hustmyer (1974), "Effects of Four Psychological Primary Colors on GSR, Heart Rate and Respiration Rate," *Perceptual and Motor Skills*, 38 (June), 763–66.

Kellaris, James J. and Robert J. Kent (1993), "An Exploratory Investigation of Responses Elicited by Music Varying in Tempo, Tonality, and Texture," *Journal of Consumer Psychology*, 2 (October), 381–401.

Kueller, Rikard and Byron Mikellides (1993), "Simulated Studies of Color, Arousal, and Comfort," in *Environmental Simulation: Research and Policy Issues*, Robert W. Marans and Daniel Stokols, eds New York: Plenum Press, 163–90.

Levy, Michael and Barton A. Weitz (2004), *Retailing Management*. New York: McGraw-Hill/ Irwin.

Mehrabian, Albert and James A. Russell (1974), *An Approach to Environmental Psychology*. Cambridge, MA: MIT Press

Milliman, Ronald E. (1982), "Using Background Music to Affect the Behavior of Supermarket Shoppers," *Journal of Marketing*, 46 (Summer), 86–91.

Novak, Thomas P., Donna L. Hoffman, and Yiu-Fai Yung (2000), "Measuring the Customer Experience in Online Environments: A Structural Modeling Approach," *Marketing Science*, 19 (Winter), 22–42.

Prior, Molly (2001), "Toys 'R' Us Times Sq. Store Raises Retailtainment Bar," *DSN Retailing Today*, 40 (December 10), 1, 46.

Reynolds, Kristy E. and Sharon E. Beatty (1999), "A Relationship Customer Typology," *Journal of Retailing*, 75 (Winter), 509–523.

Russell, James A. (1978), "Evidence of Convergent Validity on the Dimensions of Affect," *Journal of Personality and Social Psychology*, 36 (October), 1152–68.

——— and Lisa Feldman Barrett (1999), "Core Affect, Prototypical Emotional Episodes, and Other Things Called Emotion: Dissecting the Elephant," *Journal of Personality and Social Psychology*, 76 (May), 805–819.

——— and Geraldine Pratt (1980), "A Description of the Affective Quality Attributed to Environments," *Journal of Personality and Social Psychology*, 38 (February), 311–22.

Sherman, Elaine, Anil Mathur, and Ruth B. Smith (1997), "Store Environment and Consumer Purchase Behavior: Mediating Role of Consumer Emotions," *Psychology & Marketing*, 14 (July), 361–78.

Smith, Patricia C. and Ross Curnow (1966), "'Arousal Hypothesis' and the Effects of Music on Purchasing Behavior," *Journal of Applied Psychology*, 50 (June), 255–56.

Spies, Kordelia, Friedrich Hesse, and Kerstin Loesch (1997), "Store Atmosphere, Mood and Purchasing Behavior," *International Journal of Research in Marketing*, 14 (February), 1–17.

Stone, Gregory P. (1954), "City Shoppers and Urban Identification: Observations on the Social Psychology of City Life," *American Journal of Sociology*, 60 (July), 36–45.

Sweeney, Jillian C. and Fiona Wyber (2002), "The Role of Cognitions and Emotions in the Music-Approach-Avoidance Behavior Relationship," *Journal of Services Marketing*, 16 (March), 51–69.

Tauber, Edward M. (1972), "Why Do People Shop?" *Journal of Marketing*, 36 (October), 46–49.

Tedesco, Bruce G. (2000), "Web Sites Need Advice on Flow," *Marketing News*, 34 (August 14), H26.

Thayer, Robert E. (1970), "Activation States as Assessed by Verbal Report and Four Psychophysiological Variables," *Psychophysiology*, 7 (July), 86–94.

Valdez, Patricia and Albert Mehrabian (1994), "Effects of Color on Emotions," *Journal of Experimental Psychology: General*, 123 (December), 394–409.

Westbrook, Robert A. and William C. Black (1985), "A Motivation-Based Shopper Typology," *Journal of Retailing*, 61 (Spring), 78–103.

Wirtz, Jochen and John E.G. Bateson (1999), "Consumer Satisfaction with Services: Integrating the Environment Perspective in Services Marketing into the Traditional Disconfirmation Paradigm," *Journal of Business Research*, 44 (January), 55–66.

Yalch, Richard F. and Eric R. Spangenberg (2000), "The Effects of Music in a Retail Setting on Real and Perceived Shopping Times," *Journal of Business Research*, 49 (August), 139–47.

Yoo, Changjo, Jonghee Park, and Deborah J. MacInnis (1998), "Effects of Store Characteristics and In-Store Emotional Experiences on Store Attitude," *Journal of Business Research*, 42 (July), 253–63.

U.S. Postal Service Statement of Ownership, Management, and Circulation (Required by 39 U.S.C. 3685)

1. Title of Publication. *Journal of Marketing*
2. Publication No. 0022-2429
3. Filing Date. October 1, 2005
4. Issue Frequency: Quarterly
5. No of Issues Published Annually 4
6. Annual Subscription Price. $90
7. Complete Mailing Address of Known Office of Publication: 311 South Wacker Dr., Suite 5800, Chicago, IL 60606-2266
8. Complete Mailing Address of the Headquarters or General Business Offices of the Publisher: 311 South Wacker Dr., Suite 5800, Chicago, IL 60606-2266
9. Full Names and Complete Mailing Address of Publisher, Editor, and Managing Editor: Publisher Jack Hollfelder, 311 S Wacker Dr., Suite 5800, Chicago, IL 60606-2266; Editor Ruth N Bolton, W P Carey Schl of Business, Arizona State Univ, Main Campus, BAC 460, Tempe, AZ 85287-4106; Managing Editor: Francesca V Cooley, 311 S Wacker Dr., Suite 5800, Chicago, IL 60606-2266
10. Owner: American Marketing Association, 311 South Wacker Dr, Suite 5800, Chicago, IL 60606-2266
11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1% or More of Total Amount of Bonds, Mortgages, or Other Securities (If there are none, so state): None
12. For completion by nonprofit organizations authorized to mail at special rates The purpose, function,and nonprofit status of this organization and the exempt status for federal income tax purposes Has Not Changed During Preceding 12 Months
13. Publication Name. *Journal of Marketing*
14. Issue Date for Circulation Data Below October, 2005
15. Extent and Nature of Circulation:

	Average no. copies each issue during preceding 12 months	**Actual no. copies of single issue published nearest to filing date**
a. Total No Copies (Net Press Run)	9,963	11,050
b. Paid and/or Requested Circulation		
(1) Paid/Requested Outside-County Mail Subscriptions	8,624	8,634
(2) Paid/Requested In-County Mail Subscriptions	0	0
(3) Sales Through Dealers and Carriers, Street Vendors, Counter Sales & Other Non-USPS Paid Distribution	375	1,500
(4) Other Classes Mailed Through the USPS	0	0
c. Total Paid and/or Requested Circulation *[Sum of 15b (1), (2), (3) and (4)]*	8,999	10,134
d. Free Distribution by Mail *(Samples, Complimentary, and Other Free) [Sum of 15d (1), (2) and (3)]*	11	11
e. Free Distribution Outside the Mail (Carriers or Other Means)	0	0
f. Total Free Distribution *(Sum of 15d and 15e)*	11	11
g. Total Distribution *(Sum of 15c and 15f)*	9,010	10,145
h. Copies Not Distributed	953	905
i. Total *(Sum of 15g and h)*	9,963	11,050
j. Percent Paid and/or Requested Circulation *(15c/15g x 100)*	99.9% paid circulation	99.9% paid circulation

16. This Statement of Ownership will be printed in the January, 2006, issue of this publication
17. Signature and title of editor, publisher, business manager, or owner Jack Hollfelder, Publisher. I certify that the information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions *(including fines and imprisonment)* and/or civil sanctions *(including multiple damages and civil penalties)*

Wolfgang Ulaga & Andreas Eggert

Value-Based Differentiation in Business Relationships: Gaining and Sustaining Key Supplier Status

Many business customers today consolidate their supply bases and implement preferred supplier programs. Consequently, vendors increasingly face the alternative of either gaining a key supplier status with their customers or being pushed into the role of a backup supplier. As product and price become less important differentiators, suppliers of routinely purchased products search for new ways to differentiate themselves in a buyer–seller relationship. This research investigates avenues for differentiation through value creation in business-to-business relationships. The results suggest that relationship benefits display a stronger potential for differentiation in key supplier relationships than cost considerations. The authors identify service support and personal interaction as core differentiators, followed by a supplier's know-how and its ability to improve a customer's time to market. Product quality and delivery performance, along with acquisition costs and operation costs, display a moderate potential to help a firm gain and maintain key supplier status. Finally, price shows the weakest potential for differentiation.

Over the past decades, there has been a growing recognition among scholars and practitioners that collaborative buyer–seller relationships represent a source of competitive advantage. The marketing literature began to acknowledge the need to manage interfirm relationships as strategic assets in the 1980s (Jackson 1985; Webster 1984). Since then, numerous empirical studies have investigated the benefits of close buyer–seller relationships and their potential to achieve superior results (Cannon and Homburg 2001; Ganesan 1994; Hewitt, Money, and Sharma 2002; Jap 1999; Kalwani and Narayandas 1995; Lyons, Krachenberg, and Henke 1990).

Along with the recognition of the strategic importance of supplier relationships, firms have fundamentally changed the way they manage supplier portfolios, and customers have increasingly moved away from an adversarial relationship management style with many suppliers toward a logic of building long-term relationships with selected key suppliers (Kalwani and Narayandas 1995; Narayandas and Rangan 2004; Spekman 1988). Consequently, many firms have consolidated their supply bases. For example, at the end of 2003, Sony Corporation announced Transformation 60, a global program intended to boost operating profit margins from 4% to 10% over a period of three years. In a strong move of reforms, the group planned, among other measures, to reduce the number of suppliers for components and raw material from 4700 to 1000. Over the same period, the number of registered parts was to be reduced from 840,000 in 2003 to 100,000 to achieve cost reductions and quality improvements.

Such a profound shift in the management of buyer–seller relationships raises several challenges for customers and suppliers alike. From a customer perspective, firms need to differentiate among qualified suppliers and identify vendors that will form their consolidated supply base. This is particularly relevant for key components, for which single sourcing or multiple single sourcing tend to become the norm rather than the exception. Consequently, customers rely on performance evaluation tools and vendor development programs to strengthen their ties with preferred suppliers.

In turn, with growing supply base consolidation, many suppliers are challenged to move into a main supplier position and to defend this top spot against competitors. In short, vendors face the alternative of either gaining key supplier status with their customers or being pushed into the role of a backup supplier, capturing only a small share of a customer's business. This challenge is further reinforced by the trend toward product commoditization in many markets (Rangan and Bowman 1992). As product and price become less important differentiators, suppliers of routinely purchased products search for new ways to differentiate themselves through improved customer interactions (Vandenbosch and Dawar 2002).

From an academic point of view, differentiation in business relationships can be researched from a value-based perspective. To be effective, differentiation must contribute to customer value by either providing benefits to the cus-

---

Wolfgang Ulaga is Associate Professor of Marketing, Marketing Department, ESCP-EAP European School of Management (e-mail: wulaga@escp-eap.net). Andreas Eggert is Professor of Marketing, University of Paderborn (e-mail: andreas.eggert@notes.upb.de). This research was conducted while the first author was a visiting associate professor in the Mendoza College of Business at the University of Notre Dame. The authors thank Joe Guiltinan and the Marketing Department at Notre Dame for their support in this project. In addition, the authors thank the Institute for the Study of Business Markets at the Smeal College of Business Administration, Pennsylvania State University, for providing financial support for this study. The authors also acknowledge the technical assistance provided by the Institute for Supply Management in collecting the data for this research. Finally, the authors extend their gratitude to Joe Cannon and the anonymous *JM* reviewers for their constructive comments.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

tomer or lowering a customer's costs. The key role of customer value has been well established in the business marketing discipline over the past decade (Anderson, Jain, and Chintagunta 1993; Anderson and Narus 1998, 2004). Building on this literature, scholars have advanced the understanding of relationship value in recent years, both conceptually and empirically (Anderson 1995; Hogan and Armstrong 2002; Krapfel, Salmond, and Spekman 1991; Walter et al. 2003; Wilson 1995; Wilson and Jantrania 1995). However, a review of the emerging literature on relationship value reveals several unresolved research issues.

From a conceptual perspective, a sound understanding of the dimensions that drive value creation in manufacturer–supplier relationships is needed. Previous studies explored either individual drivers or subsets of value-creating dimensions in business relationships. For example, Cannon and Homburg (2001) investigate value creation through cost reduction in a manufacturer–supplier relationship. However, a thorough investigation of value drivers that integrate the dimensions into a framework grounded in managerial practice is still missing.

From a measurement perspective, guidelines need to be provided for modeling relationship value as an overall measure (Parasuraman 1997). Previous research does not provide any guidance as to how the different dimensions should be consolidated. Should a reflective or formative measurement approach be used (Diamantopoulos and Winklhofer 2001)? Indeed, researchers have conceptualized value as a reflective construct without justifying their approach (Lapierre 2000). More research is needed in this area.

Finally, from a managerial perspective, there needs to be a greater understanding of which value-creating dimensions of a business relationship represent a promising basis for differentiation. In other words, from a vendor perspective, which value drivers are key when competing for main supplier status? Which dimensions show only a limited potential for differentiation and are to be considered a "must-have"? In turn, from a customer perspective, which dimensions help discriminate among competing vendors and further consolidate a supply base?

To address these challenges, our article is structured as follows: Building on Cannon and Homburg's (2001) categorization of relationship costs, we develop an integrated framework of relationship value. Next, we present the results of our qualitative study aimed at identifying and describing key value drivers in each category of our framework. We then present the results of a nationwide cross-sectional survey among senior purchasing managers. Finally, we discuss the implications of our research for both academics and managers and conclude by outlining study limitations and an agenda for further research.

Toward an Integrated Framework of Relationship Value

Creating superior customer value is key to a company's long-term survival and success (Slater 1997; Woodruff 1997). In business markets in particular, customer value is regarded as the cornerstone of the marketing management process (Anderson and Narus 2004). Yet despite its importance, research on customer value in business markets is still in an early stage (Flint, Woodruff, and Gardial 2002). Although value assessment studies enjoy a long tradition in business marketing, they typically focus on the value of the physical product, neglecting relational dimensions of customer-perceived value (Dwyer and Tanner 1999).

In recent years, researchers have adopted a relational approach that investigates customer value from a relationship marketing perspective (Payne and Holt 1999). It is clear that the value of a business relationship is a multidimensional concept that reaches beyond the price versus quality trade-off that is prevalent in consumer research (Gassenheimer, Houston, and Davis 1998). To date, however, there is a lack of concurrence in the marketing literature on the salient value dimensions of a business relationship (Woodall 2003).

Recently, Cannon and Homburg (2001, p. 29) directed attention toward the costs incurred in business relationships, indicating that "one method for creating value is to reduce costs in commercial exchange." On the basis of a multidisciplinary literature review, they identified three sources of relationship costs: (1) direct costs, (2) acquisition costs, and (3) operations costs. They further suggested that a supplier's success in lowering a customer firm's cost in each of the three categories leads the customer to expand its business with the supplier. In discussing directions for further research, Cannon and Homburg recognized that "cost reduction is only one source of value in business relationships. A more comprehensive theory would consider costs and benefits beyond economic costs."

Following this line of reasoning, we suggest thinking of relationship value in terms of a 2 × 3 matrix that distinguishes between two fundamental dimensions of value creation (i.e., benefits and costs) and three levels at which these drivers operate (i.e., the core offering, the sourcing process, and the customer firm's internal operations). Rather than relying on a hypothetico-deductive approach, we defined the value drivers from a managerial perspective (see, e.g., Bendapudi and Leone 2002; Flint, Woodruff, and Gardial 2002). We conducted depth interviews with purchasing managers using a grounded theory approach (Glaser and Strauss 1967). We then used the results of this qualitative study as a basis for a cross-sectional survey to investigate the various value drivers' potential of serving as key differentiators and to develop an overall measure of value creation in key supplier relationships.

Qualitative Study

Data Collection and Analyses

Sampling procedure. We gathered data through depth interviews with purchasing managers in manufacturing companies located in the Midwest of the United States. Participants were contacted through the local chapter of the Institute of Supply Management (ISM) and through the alumni network of a midwestern university. The sampling process ceased when saturation was reached, which was indicated by information redundancy. Our final sample consisted of ten participants in nine manufacturing organizations. This is consistent with sample sizes recommended by scholars for exploratory research purposes (McCracken 1988, p. 17).

Sample characteristics. Table 1 displays key sample characteristics. Our sample consisted of manufacturers in a variety of areas, such as aircraft landing systems, amplifiers and microphones, audiovisual video projection equipment, automobiles, household appliances, orthopedic implants, and vacuum pumps. Components considered by our study's participants also varied significantly. For example, customers purchased aluminium wheel forgings, car seats, electronic components, motors, springs, and surgical instruments. The size of participating companies ranged from small and medium-sized manufacturers to multinational corporations, and the selected buyer–supplier relationships had been in place between 2 and 25 years.

In developing our sample, we aimed to maximize diversity among participants to discover various potential value drivers and their underlying facets. At the same time, study participants and firms also needed to share several characteristics to allow for comparability. For example, because we relied on key informants, it was critical to select managers who were influential decision makers involved in selecting and monitoring supplier relationships. Consequently, we invited only senior-level managers to participate in the study. In addition, key manufacturing characteristics of companies needed to allow for comparability across firms. Thus, only companies in comparable industries were contacted (Standard Industrial Classification codes 34–38).

Interview guide. Our interview guide was composed of three sections. In the first section, we asked respondents to select a specific component they purchased from at least

TABLE 1
Qualitative Study Sample

Name	Participant Background	Company Activity and Size	Component Purchased and Supplier Context
Denice	Contract Admin. Supervisor, age 42, 22 years in purchasing	Amplifiers Sales: $75 million Employees: 450	Electronic components One main supplier, multiple other suppliers
Frank	Manager, e-procurement, age 35, 12 years in industry, 4 years in purchasing	Automobiles Sales: $177 billion Employees: 347,700	Car interior system One system supplier, four alternative suppliers
Jack	Director of Purchasing, age 45, 25 years in production, quality control, and purchasing	Aircraft landing systems Sales: $9.7 billion Employees: 37,500	Aluminium forgings One main supplier (90%), one backup supplier (10%)
Jeff	Senior Purchasing Supervisor, age 38, 13 years in purchasing	Automotive brakes Sales: $6.2 billion Employees: 21,750	Pins One main supplier (95%), multiple other suppliers
Jerry	Purchasing Manager, age 50, 19 years in planning and inventory control, 9 years in purchasing	Video screens Sales: $110 million Employees: 500	Springs One main supplier, multiple other suppliers
John	Purchasing Manager, age 38, 3 years in production planning, 16 years in purchasing	Electric sensors Sales: $458 million Employees. 5313	Circuit Boards One main supplier (80%), one backup supplier (20%)
Mary	Purchasing Manager, age 45, 25 years in quality control and purchasing	Vacuum pumps Sales: $82 million Employees: 400	Electric motors Two main suppliers, multiple other suppliers
Richard	Director, Strategic Sourcing, age 40, 18 years in purchasing	Orthopedic implants Sales: $1.2 billion Employees: 3600	Surgical instruments Two main suppliers, multiple other suppliers
Scott	Director Global Commodities, age 38, 6 years in engineering, 7 years in purchasing	Household appliances Sales: $11 billion Employees: 59,408	Electric motors Two main suppliers, multiple other suppliers
Shawn	Project Leader, Global Sourcing, age 28, 4 years in purchasing	Aircraft landing systems Sales: $9.7 billion Employees: 37,500	Aluminium forgings One main supplier (90%), one backup supplier (10%)

Notes: Names are pseudonyms All participants are key decision makers in the purchasing departments of their firms.

two suppliers. We further invited participants to describe the product, how it was used in the manufacturing process, its relative importance in the final product, and the supply market. Customers needed to maintain a collaborative relationship with suppliers for the selected product. Therefore, we used characteristics described in the extant literature (Jap 1999) to ask participants why they considered a particular relationship collaborative. The purpose of this initial stage was to ensure that respondents considered a specific use situation and were prepared to compare alternative buyer–supplier relationships.

In the second part, we asked respondents to describe how suppliers created value for their organizations and to provide examples from the specific supplier relationship under consideration. To facilitate the process, participants were asked to describe activities between the supplier and the manufacturer, which in turn allowed the interviewer to probe the different benefit and cost dimensions perceived in the relationship. Particular attention was given to the comparison of each company's main supplier and its second-best alternative supplier of the same component. In this section, we carefully phrased our questions in such a way that respondents' testimonies were elicited in an unobtrusive and nondirective manner to avoid potential pitfalls of "active listening" (McCracken 1988, p. 21). Thus, the objective in this stage was to facilitate the emergence of relationship value categories grounded in the managers' own language rather than to capture merely the dimensions previously specified by the investigators. Finally, in the third part of the interview guide, we invited participants to describe themselves and their company background.

Analysis and interpretation. On average, interviews lasted one and a half to two hours. Each interview was audiotaped and transcribed verbatim. We used grounded theory coding (i.e., open, axial, and selective coding) to identify relationship value drivers and to describe their facets (see, e.g., Flint, Woodruff, and Gardial 2002). We assessed the trustworthiness of our findings by applying the techniques of triangulation, informant feedback, and replication (Miles and Huberman 1994; Strauss and Corbin 1998). To improve content validity further, we presented the study's methodology and findings during a workshop with 27 purchasing professionals at ISM. We asked participants to assess the descriptions of the core value drivers and to comment on how well they reflected their practice. Managers largely agreed on the meaning of the value drivers. Therefore, we made only a few changes in wording and illustrations after this step.

Field Study Findings

Key supplier relationships. Across the board, study participants mentioned that their companies had reduced their supply bases over time, as Jack illustrated:

> Over the last ten years, we have been streamlining our supply base. We used to have over 300 suppliers. We now have driven that down to 115 elite suppliers that do a great job for us.

Along with the trend toward downsizing supplier portfolios, the remaining vendors captured larger shares of purchasing volumes. Consider the following comment from John:

> We have between 60 and 65 suppliers of production items, and 10 or 11 of those make up our key suppliers, that is, 80% of the dollars we spent. For those suppliers, we have a specific supplier development program.

Participants identified situations of sourcing key components from at least two firms, that is, the main supplier and its challenger. The following quotation from Richard illustrates this:

> We have about 60 suppliers of surgical instruments, and we spend $63 million on them. The number one supplier has $10 million of it, and the number two supplier has $7 million. These are the two suppliers that stick out. It then drops dramatically. The other suppliers will get down to a couple of $100,000 each.

After participants described their supply base and selected a specific supplier relationship, we moved on to explore how vendors gained and sustained key supplier status with a customer. As we show in Table 2, nine value drivers emerged from our depth interviews. We classified these drivers by value dimension (i.e., relationship benefits and costs) and value sources (i.e., value creation through the core offering, within the sourcing process, and at the level of the customer's operations). In the next section, we present these core value-creating dimensions.

Value Creation Through the Core Offering

We asked managers what was needed for a vendor to gain a key supplier status. Participants consistently described a set

TABLE 2
Value Drivers in Key Supplier Relationships

	Relationship Value Dimensions	
Sources of Value Creation	**Benefits**	**Costs**
Core offering	Product quality Delivery performance	Direct costs
Sourcing process	Service support Personal interaction	Acquisition costs
Customer operations	Supplier know-how Time to market	Operation costs

of core dimensions. They identified product quality and delivery performance as the two benefit dimensions at this level of value creation. The corresponding relationship cost dimension was direct product costs. Respondents explained that vendors needed to meet these must-haves. Therefore, we named this first level of value creation "the vendor's core offering." Jerry illustrated this point:

> I think you have to look at their quality, delivery performance, and price, and then you need to evaluate all of that. If anyone of those is weak and you cannot get them to improve, then you have to change. All of those factors have to be in harmony.... I would say that is the ground zero. That is where they need to be in order to be considered.

Product quality. Study participants invariably emphasized performance and reliability in a supplier's product offering. In addition, respondents mentioned the importance of meeting the customer's technical specifications, as the following comment from John illustrates:

> For a circuit board, we look for a specific output. We want to make sure there are no cracks in the hybrids. None of the parts must be skewed on the circuit board. The part has to be dimensionally and electrically correct per a set of specifications. We gave the supplier a drawing that outlines the physical dimensions, the distances between the pads, where the capacitors are supposed to be, [and] where the traces are, and then there are the electrical specifications the supplier has to meet—that is, the output this circuit board has to produce.

Respondents also mentioned the importance of delivering consistent quality levels over time. Notably, a vendor's key supplier status may shelter the firm from short-term sanctions. For example, respondents mentioned that their firms worked with key suppliers to solve a vendor's quality problems through joint teams. Consider the following statement from Jeff:

> We work with a supplier that doesn't hit the [quality] target.... We go in and find out what problems they have. Our Supplier Quality Group actually helps them to develop test procedures, to find the latest technology, [and] to help them fix issues on the line. Quality is a major component, but we do help them out in identifying where they can improve.

Several participants in our study underscored the pressure for continuous improvements. For example, Scott illustrated the need for cutting-edge product quality:

> One of the key elements of a supplier relationship is having the best-in-class quality for the components that they supply.... We expect suppliers to work with us and try to continuously drive quality. As you see more European and Asian products come into our country, the bar gets raised all the time, and you have to benchmark yourself against different competitors than those you had in the past.

In speaking about quality's contribution to value creation in a supplier relationship, participants indicated that superior product quality levels may indeed serve to some extent as a differentiator among suppliers. Consider the following comment from Frank:

> They [market research firm] do a customer satisfaction survey every year on vehicles. The Japanese have dominated that, but we have made substantial improvements in the last four years One of the reasons is that we have been sourcing with higher-quality suppliers.... When we are sourcing with [first-tier supplier], we can go through all of their manufacturing locations and see nothing but great marks for them. That is one way a supplier can charge a little more. It's because they have a proven track record to supply excellent parts.

However, it also became apparent that it is increasingly difficult for suppliers to compete on quality only. As Frank said,

> Most of our suppliers are all delivering a quality product. Our quality group will work with suppliers to make sure that they are meeting our quality objectives. So then we have a bid list which the buyer can take.

On the basis of these findings, we identify product quality as the extent to which the supplier's product meets customer specifications. Key quality aspects are performance, reliability, and consistency over time. Our study confirmed the potential for differentiation through superior quality. Yet its role appeared to be limited.

Delivery performance. Purchasing managers identified three value-creating aspects in this area. First, suppliers create value by consistently meeting delivery schedules (on-time delivery). Frank talked about consequences of late deliveries in the car industry:

> The most important is that they are on time with the due date that we have on the purchase order. We try to streamline how much inventory we have as much as possible. If the suppliers are not meeting their schedules in a timely fashion, that causes a big hiccup and may result in premium freight to get parts here.

According to the respondents, a second way to create value is to adjust to changes in delivery schedules due to spikes in demand or changes in the product mix (delivery flexibility). Managers valued suppliers' responsiveness for emergency deliveries, as this quote from Jack illustrates:

> You will notice the suppliers that are your best friends. When you are down because someone didn't count the parts right, you call them up.... A lot of suppliers will turn their shop around to keep us running. Those are the things that really stick in your mind. Suppliers that go above and beyond what a typical supplier will do.

As a third source of value creation in this area, participants mentioned suppliers' capacity to deliver the right parts consistently (accuracy of delivery). Minimizing missing or wrong parts in shipments saves time and effort for the customer.

In summary, we identified delivery as a second value driver. Suppliers add value by consistently meeting delivery schedules (on-time delivery), by adjusting to delivery changes (flexibility), and through their capacity to deliver the right parts consistently (accuracy).

Direct product costs. There was remarkable commonality across respondents when they discussed the role of direct product costs. Participants identified the actual price charged by a supplier as the core relationship cost driver. Managers focused on "average," "fair," or "reasonable" market price. As Jerry explained,

> We don't want any supplier that loses money, but we want them to be tough too. We want them to give us the best price as they can at the margin they can live with. They need to make money as well as we do.

On further probing into the topic, participants explained that they were hesitant to overemphasize price considerations. For example, Scott said,

> There is always somebody cheaper, but it is one thing to say you can make a part at a lower cost, and it is another to say you can do both: Support our business needs *and* supply us at a lower cost. We are very hesitant to be over-aggressive. There are customers who wave these prices in front of suppliers that come from people who can't make or produce. That is a dangerous game to get into with your supply base.

In addition, participants' firms face a strong pressure to reduce costs in their respective markets. Suppliers are expected to commit to annual price reductions within long-term agreements and to pass cost savings on to customers. In turn, respondents agree to higher order volumes to compensate for lower prices. As John commented,

> We tell our suppliers that we want 4% of cost reductions every year. That is pretty typical in the automotive industry. So, speaking of suppliers, I want a guy who is willing to continuously improve his operations. One who finds ways to lower his costs, so that some of that can be passed on to me. So, I can say to my management, "Hey, instead of paying $1.50 for this item, now I will start paying $1.40." There is value that can be added that way. It allows us to offer our own customers a cost reduction.

Similar to our findings in the area of product quality, a key supplier status may protect a vendor from competition, at least in the short run. Customers indicated that they provided main suppliers with an opportunity to match competition. In addition, they assisted key suppliers in driving down prices through joint cost reduction programs, an initiative not granted to backup suppliers.

In summary, we learned from our interviews that managers focused primarily on price as the key relationship cost driver. Participants valued both the supplier's capacity to offer a fair market price and its commitment to reduce prices continuously. In turn, key suppliers benefited from increased order volumes and an opportunity to match competition.

Value Creation in the Sourcing Process

Beyond the core offering, vendors create value in the sourcing process. Relationship benefit drivers identified at this level are the supplier's service support and the personal interaction between both parties. The mirroring value-creating cost dimension is acquisition costs.

Service support. The supplier's capacity to provide value-added services was another common theme. The following comment from John illustrates in general terms the role of services in a buyer–seller relationship:

> The other thing would be just general service. If we have a quality problem, are they available to answer that question? I talked about it specifically on the technical side, but it is actually bigger than that.... I will call it a general service category; are they there for us when we have problems that may or may not be related to the technical design?

When probing further into the importance of services in value creation, three facets emerged from our depth interviews in this area. First, respondents highly valued a supplier's responsiveness, that is, the vendor's willingness to address customers' concerns in ongoing relationships. Shawn mentioned responsiveness as a key service element:

> I'd say one of the key drivers of differentiation has to do with responsiveness. For example, our products are very complex. We are talking about hundreds of pages of documentation, and you are not going to get it right the first time. We don't expect you to, but what we do expect is a very concerted effort to get there. One of the main things that hurt us is the fact that a supplier is unresponsive to the problems we are having. So, at the end of the day, responsiveness is a huge part of the equation.

A second facet of service support was the supplier's capacity to manage exchanges of information (information management). This subdimension had two ramifications: First, respondents mentioned that customers frequently changed their requirements. Whenever such changes occurred, suppliers were expected to follow through in a timely manner. Thus, speed of information represented an important aspect of information management. Second, our participants voiced their need to receive detailed information. The following quotation from our interview with Jeff echoes both speed and appropriateness of information:

> Details are important. If we get a request for a quote from [car manufacturer] and our supplier just gives us a number on a paper, it doesn't help us out.... We get requests for changes from our customers all the time, and sometimes we need to get back to them within a few days. So speed is certainly a factor.

The third component of service support is the outsourcing of activities to suppliers. Respondents cited examples of situations in which suppliers added value by taking responsibility for particular customer activities. In our interviews, we learned that key suppliers solidified their competitive position by performing tasks on behalf of their customers. Shawn talked about how a supplier performed routine testing for his firm:

> There are several things that are service related. One of them is testing and qualification. Take, for example, aerospace grade bolts—the types of things that hold wheels together on aircrafts. This supplier will come out and do testing for you. Because of the amount of tests we have, there is not enough capacity for testing [at our plant]. So they would do the testing for us, report the results, and work with us to make improvements on the parts, and that is a value-added service.

Customers benefited from outsourcing in numerous ways. For example, by delivering integrated systems as opposed to single parts, suppliers contributed to consolidating the supply base. Consider the following comments from Frank about outsourcing:

> Consolidation in the supply base is one way in which the suppliers are creating value. For the [car model], we looked at suppliers that could provide an entire interior, so this meant that they would be supplying the instrument

> panel, the floor consol, the overhead system, the door panels, the side wall trim, garnish, the rear shelf and even the carpeting.... Suppliers take on some of the subassembly operations. Their facilities are close to our assembly plants, which leads to large reductions in plant space. So we can either build more vehicles or reduce the time that it takes to build a vehicle.

In summary, we identified service support as an important relationship value dimension. In addition to the supplier's level of responsiveness, respondents mentioned that key suppliers create value through information management and outsourcing of activities.

Personal interaction. Managers noted numerous examples of value creation through personal interaction. They held the development of relationships at an individual level in high regard. Consider the following statement from Denice:

> I think the personal issues are important, at least for our company, because there is such a close interaction between their folks and ours. If you have people that for one reason or another do not get along, it strains the relationship and is counterproductive. So I believe that the value of the interpersonal skills is quite high.

Developing interpersonal ties improved problem solving and communication and led to a better understanding of each partner's goals. These benefits were regarded as contributing to the growth of a relationship as a whole. As Jeff commented,

> When I first came to [company], the person who I took the position over [from] was not well liked by the supplier. They voiced their opinion as such. I changed the way we work with them. When dealing with these people, I look at them as more than just a tool. I look at them as a person. To me, it is very important. If you don't have that, I don't think you can function well in the industry. I think it [the relationship] grew more after I came on board because of the relationship that I developed with them. I think they opened up and understood what they need[ed] to do to move the relationship ahead.

Managers considered suppliers' key contact personnel important facilitators for conducting business. For example, when comparing two suppliers of electrical motors, Mary mentioned that one company was definitely much easier to work with than the other:

> One company is much easier to do business with than the other, just because they give us their names and e-mail addresses. The quality people know the quality managers; they can call them up and talk to them. All of those lines are very open. The other company, well, you have to do all of those things through the salesperson. It is much more difficult to get the answers because now, my quality person is calling me asking a detailed question, and I have to call the salesperson and give him a detailed question, and he goes and talks to his people, who ask a question that neither one of us knows an answer to.

Finally, respondents indicated that personal interactions should be developed at all levels of the organizations. In particular, strong involvement of the vendor's top management was viewed as an indicator of a good working relationship. As Richard illustrated,

> One of the things that are of value is that you have the presidents of these companies show their face here. We like the comfort of having the president here and knowing what is going on.... I can call the president of a company and get him to shake the stick over there. It's the name, the face: "You can count on me if you need to use the silver bullet to get something done." I think that is very important.

Although participants viewed personal interaction as an important asset, they also cautioned against its potential pitfalls. For example, Richard referred to his company's internal rules of conduct with respect to handling gratuities:

> As a buyer, you really have to be careful of how you are perceived within the organization. If you are walking out of the front door every day to the supplier's car to get lunch, that is an issue. We don't mind the occasional, but for the most part, we have a very low tolerance level for it. The individuals that work for us are well-paid individuals. We don't need to take advantage of the other stuff.

In summary, personal interaction emerged as an important value driver. Knowing the supplier's key contact personnel, getting along well with the vendor's representatives, and involving a supplier's top management all contribute to high levels of personal interaction.

Acquisition costs. Participants highly valued a supplier's willingness to take costs out of the sourcing process. Inventory management emerged as a first opportunity for cost reductions in this area. For example, Scott mentioned how having suppliers manage customers' inventories improved cash flows:

> Inventory management is of tremendous value to us because we are able to take costs off our books and reinvest in other parts of our business to become more efficient. Suppliers that really work well with us move to a "pay-and-consumption" model. We work with them to reduce inventories by trying to get visibility throughout the whole supply chain.

Order handling represented a second way to reduce acquisition costs. Participants' firms had standardized order processes and had downsized staff. For example, Jack mentioned that he used to have 14 buyers in his purchasing department. Over a period of 20 years, he was down to 6 buyers. Thus, participants relied more heavily on suppliers for order handling. As John commented,

> We've set up a Kanban system. So I never, ever issue a purchase order for the delivery of those circuit boards. The operator on the line knows that when he gets down to one box, he sends a notification to the supplier, "Hey, Mister Supplier, I am down to one box, send me two more boxes" So he does that, and I never see the order for that.

Inspection costs represented a third vector of value creation through reductions of acquisition costs. Customers reduced incoming inspections as suppliers consistently met expectations. John talked about how these measures translated into cost savings:

> We track quality each month through our no-incoming-verification program. If a supplier's product meets our quality criteria so many months, we don't inspect the

> product anymore here. As soon as it comes in the door, it goes to the assembly line.

In summary, suppliers find multiple ways of adding value by taking costs out of the sourcing process. Drivers of cost reductions identified in this area were inventory costs, order-handling costs, and incoming-product inspection costs.

Value Creation in Customer Operations

Customer operations represent another domain of value creation. The benefit dimensions identified in this area are the supplier's specific know-how and its capacity to improve time to market. The corresponding cost dimension is operations costs.

Supplier know-how. A vendor's know-how emerged as another common theme in our interviews. Participants provided many examples of situations in which access to a supplier's technical expertise was viewed as highly valuable. Customer organizations benefited from their suppliers' know-how in several ways. First, participants mentioned that vendors' extant knowledge of supply markets provided them with an opportunity to present customers with new sourcing alternatives. The following comment from Denice echoes this idea:

> [This supplier] comes to our facility at least once a week and works very extensively with our design engineers. So if they have an application for a semiconductor, they give him the performance specifications and let him come back with one or two solutions. This reduces the investment we have to make in terms of knowledge and experience of individual components. The semiconductor market is changing every single day, and to keep abreast of all new products in the market would be very expensive for us. So we rely very heavily on this supplier to bring us those products and give us awareness. It also means that we can have younger people on the engineering staff. They don't have the same background.

Second, a supplier's thorough understanding of a customer's operations and a long-standing experience with a client's products both created opportunities for vendors to add value in the improvement of existing products. As Shawn explained,

> Suppliers that have a lot of experience with your products know how to make parts efficiently and effectively. A supplier that has done a specific part for years can turn around and make a change at a third of the cost of a new supplier. Their lead time is usually less also because they can do the setup in the dark, and they make suggestions.... The expertise in a relationship is amazing in terms of the value it creates. Their experience and knowledge really protect them from global competition.

Finally, suppliers assist customers in the development of new products. Key suppliers are frequently involved early on in suggesting innovative solutions and taking costs out of a new product up front. In speaking about early involvement of key suppliers in the new product development process, Richard commented,

> The knowledge base that suppliers build up is one of the key characteristics that we look for in a supplier. For example, we don't manufacture our own surgical instruments. So we don't have the knowledge base of instrument makers walking around our shop floor. If we design a new product, we rely on the expertise of the instrument makers. They need to come back to us with their experience and say, "Your design can't be manufactured. It won't work, or it will cost you a lot of money." We involve them early on in the process. They need to tell us where we need to put costs into the instrument and where it is not necessary. That goes back to the collaboration in our preproduction planning meeting. The supplier is there, our development engineer is there, and they can talk back and forth

In summary, a supplier's know-how provides many opportunities to add value in a manufacturer–supplier relationship. Key sources of value creation are the vendor's deep knowledge of the supply market, prior experience with customer operations and products, and an early involvement in new product development. We also learned that key suppliers benefit from their preferred status because the interactions between key supplier status and know-how are mutually reinforcing. Indeed, holding a specific expertise helps a supplier solidify its position. In turn, a stronger position enables a vendor to accumulate more experience with a client's products and gain better insights into a customer's operations than any other supplier.

Time to market. In our interviews, time to market emerged as an important value driver in supplier relationships. For example, in talking about the orthopedics industry, Richard highly valued a supplier's contribution to reduced cycle times:

> In our industry, you've got to be quick. If you get your product out to the market first, you win. The number of small orthopedic implants manufacturers is growing, and the really small ones are taking little pieces of market share away from us. If you want to make the big dollars in the orthopedic world, you have to get your product out there first, and right now, one of our main goals is to look at how you reduce the time it takes from the idea generation of a product to the time you get it up to market. In the past, it took us three years. The goal is to take 25% out of that.

Participants explained that their firms faced an increasing pressure to develop products faster. For example, Frank mentioned that his firm had reduced cycle times from approximately 40 months to fewer than 18 months over a period of several years. Respondents turned to suppliers in different areas to reduce cycle times. Speed of executing design work emerged as a first area, as Frank illustrated:

> Our global suppliers can get much faster turnaround in completing or changing designs. They take in a requirement and send it to an offshore facility where they have much more of the day to work with it. If we decide at 3:00 P.M. to do something, they can send it to a new facility, and by the time you come back the next morning, you already get a result.

Participants also referred to the pressure of developing prototypes faster. By developing a prototype exactly to the customer's specifications the first time, suppliers improved cycle time considerably. Richard illustrated this idea:

> You have to have instrument suppliers that are quick at what they do. For example, last week, we delivered two models to an instrument supplier, and they turned around prototypes within four days. That is an extreme, but typically, what we are looking for are lead times of 4 to 6 weeks. Two years ago, it would have been 12 to 16 weeks.

Finally, managers explained that suppliers performed testing and validation tasks faster than customers. In addition, by avoiding unnecessary product retesting, they were able to accelerate cycle times significantly. Consider Frank's statement:

> Suppliers add value through testing and validation. We have all kinds of validation requirements for our parts, whether it is bumper impact tests or sled tests for the interior airbags. Suppliers take over more and more of our validation, and they are able to do it a lot faster. That helps us to improve our cycle times. Some of our validation equipment is used 24 hours, 7 days a week. By going to supplier facilities, we speed up the validation process and get cost savings faster.

In summary, a supplier's ability to reduce time to market represented an important source of value creation in buyer–supplier relationships. Suppliers add value through accelerating design work, developing prototypes faster, and speeding up the testing and validation process.

Operations costs. Participants discussed several directions of value creation through cost reductions at the level of customer operations. One way to reduce operations costs was to take costs out of existing products. Frank provided the following example:

> We started a couple of years ago a formal supplier suggestion program. We are setting objectives to save a certain amount of money per year through submitting suggestions for how we can save money on a part. For example, a supplier will say, "We have to pierce five holes in this stamped part. We noticed the one hole isn't used anymore. Now, there is one less piercing operating we have to do. You can save that." Once engineering has approved it, we implement the idea. Then, they get a percent of the cost savings, and we get a percent.... It goes a long way into making your company stand out as to being proactive in helping us save costs. Our "Suppliers of the Year" are meeting those objectives. So it just makes sense that you would want to give those suppliers more business.

In addition to taking costs out of products, participants explained that valued suppliers typically contributed ideas of how to improve costs in the manufacturing process. For example, Scott discussed differences between two suppliers regarding costs associated with the transformation process from steel coil to stamped parts:

> We are buying steel from a certain supplier, and another supplier came in and offered us a lower price for that same type of steel. You have the cost to get it to the factory, which is the price of the steel and the freight cost, but you also have the impact on the yield of the steel as it is converted in your factory, and that is all the way from the coil to a painted or stamped part. So we will evaluate that, and we will try to get an understanding of whether there is any difference in the cost to do that conversion. Really, we are talking total cost of ownership here.

Finally, respondents mentioned tooling costs and warranty costs as another area of value creation through cost reductions in this field. As Frank mentioned,

> We'll ask suppliers what they charge us for warrantee and tooling. If you have a supplier that has shown very good quality and also has fairly low warrantee and tooling costs with their components, that is considered an advantage over another supplier that typically does not do well on these [costs]. That is an operations type of costs.

In summary, suppliers find multiple ways of adding value by taking costs out of customers' operations. Major opportunities for cost reductions in this area are product costs, manufacturing process costs, and tooling and warranty costs.

Relationship Value Index Construction

Our study provides valuable insights into how the different dimensions should be collapsed into an overall measure of relationship value. Whether value should be modeled as a reflective or formative measure is of particular importance. Diamantopoulos and Winklhofer (2001) argue that reflective specifications of latent variables often mistakenly prevail in the marketing literature. In reflective specifications, higher-order constructs are assumed to cause their dimensions rather than be caused by them. Consequently, dimensions are viewed as strongly correlated and interchangeable facets of the focal construct (Bollen and Lennox 1991). In contrast, formative specifications view a higher-order construct as being caused by its dimensions. From a formative perspective, the higher-order construct is defined by its dimensions, which do not need to be highly correlated with one another. According to Diamantopoulos and Winklhofer, the choice between a formative and a reflective specification should be based primarily on theoretical considerations.

In the past, researchers conceptualized relationship value as a reflective construct without justifying their approach (Lapierre 2000). However, the findings of our grounded theory study suggest a formative measurement approach. Indeed, the decision rules that Jarvis, Mackenzie, and Podsakoff (2003) developed for determining whether a construct is formative or reflective suggest the use of a composite latent construct model in our context. For example, in our study, causality flows from the measures to the construct. In addition, the value drivers do not need to be highly correlated with one another either within each of the value sources specified in our framework or between them (see Table 2). For example, regarding a supplier's core offering, a specific supplier may score high on product quality but low on delivery performance. Similarly, a supplier may excel in the sourcing process, yet the same firm may not differentiate itself in the area of customer operations. Indeed, study participants repeatedly described situations in which value drivers did not correlate with one another. For example, Shawn talked about the relationship between suppliers' quality and price levels:

> We can get really great pricing from places in Mexico and China, but when the product comes in and it has poor quality and they deliver late, that isn't going to keep my

> line running. Management really gets upset when you run out [of] a $.50 piece that you are trying to control inventory on and we can't ship a $200,000 order.

Similarly, Denice provided examples of situations in which suppliers performed high on the quality and delivery dimensions but scored low on personal interaction:

> A couple of years ago, we focused on the reduction of our supply base. We had a number of suppliers that were basically on time and had no quality issues, but because of the fact that there were interpersonal problems, we made the decision to cut them off altogether as a supplier. That was primarily because of their interpersonal skills. Their customer service person was not helpful or responsive [and] was very slow or abrasive. That is a very important part of the relationship.

Finally, Jack discussed a situation in which his firm maintained the relationship with a supplier mainly because of the vendor's high-quality products. In turn, delivery, pricing, and personal interaction were considered sources of conflict:

> This supplier has a good quality product, but when you look at the purchasing side, delivery, and their pricing,... [shakes his head]. When the salesman comes in here, I don't want to spend time with him, because he just gives me a "song and dance," and I usually end up getting into an argument with him. So when he comes in, he doesn't see me; he just goes up to engineering and gets his part on the blueprint, so that we have to buy it.... I try to build a better relationship. It hasn't been good for me, because I have been burnt too many times with those guys, and they don't seem to be changing their ways.

Therefore, depth interviews revealed that value drivers may correlate, but they do not need to correlate. Indeed, respondents explained that they frequently made trade-offs between value dimensions. Consequently, our qualitative study's findings suggest that from a methodological standpoint, a formative measurement approach should be used rather than reflective measures when relationship value is modeled as a multidimensional construct. More precisely, using Jarvis, Mackenzie, and Podsakoff's (2003) classification of second-order factor models, a Type IV model (i.e., a formative first-order, formative second-order model) should be specified when value creation is modeled in buyer–supplier relationships.

In summary, our qualitative study addressed the conceptual issues raised in this research. We identified and described value dimensions in buyer–seller relationships. In addition, we resolved the issue of how to integrate the dimensions into an overall index. Building on our findings, we define customer-perceived value in a key supplier relationship as a formative higher-order construct that represents the trade-off between the benefits and the costs perceived in the supplier's core offering, in the sourcing process, and at the level of a customer's operations, taking into consideration the available alternative supplier relationships.

Quantitative Study

We designed our quantitative study to address the methodological and managerial issues raised in this research. To this end, we conducted a nationwide survey among purchasing managers. This section describes both the methodology and the major findings of this second step.

Measure Development

We took several steps to ensure the content validity of our measures (Churchill 1979; DeVellis 2003). First, on the basis of our qualitative study, we developed definitions of each relationship value driver, and we generated a pool of 175 items that tap the various facets of each driver. We discussed both definitions and items in a workshop with purchasing managers. We asked the managers to assist us in the validation of each definition, and we invited them to identify the items that best captured the respective dimensions. As a result of this first step, we eliminated 77 items, leaving a reduced item pool of 98 items for nine value drivers.

In a second step, we submitted the remaining items for an item-sorting task to 15 marketing academics identified as experts in the areas of industrial marketing, relationship marketing, and research methodology. We asked these participants to assign the individual items to what they believed to be the "correct" value drivers. As Anderson and Gerbing (1991, p. 734) propose, we computed two indexes for each item to identify items that were difficult to assign to the corresponding facets.

Survey Instrument

Questionnaire development. The survey is composed of three parts. First, mirroring the approach we adopted in the qualitative study, we asked respondents to select a specific key component and to describe the final product for which it was sourced. We asked respondents to name their main supplier for the specific product and their second supplier in terms of purchasing volumes. The purpose of this initial stage was to ask respondents to consider a specific supplier and to prepare for a comparison of alternative buyer–supplier relationships. We asked participants to compare the main supplier with the second supplier for several reasons. First, depth interviews showed that managers typically compare these two alternatives when making value judgments. Second, respondents needed to use similar comparison standards to allow for meaningful comparisons. Finally, from a managerial perspective, we were interested in understanding how vendors achieve a main supplier position and defend it against their toughest competitor, that is, their challenger in terms of purchasing volume.

The second part contained a list of items that tap the relationship value dimensions. We first asked respondents to answer several statements with respect to the relationship between their firm and the main supplier. Subsequently, we asked respondents to compare their main supplier with the second supplier. All items used seven-point rating scales (1 = "strongly disagree" and 7 = "strongly agree").

In the third part of the questionnaire, we invited participants to respond to a set of questions that described themselves, their company, and the supplier relationship. Because we relied on the perceptions of key informants, we needed to ensure that respondents were competent to report on the supplier relationship. Therefore, respondents needed to fulfill minimum requirements with respect to their position, their tenure with the company, and the length of the relationship. In addition, we asked respondents for a self-assessment of their ability to portray the supplier relationship accurately. Specifically, we asked them to indicate how confident they were in answering the questionnaire, how involved with and knowledgeable they were about the supplier, and to what extent they could influence purchasing decisions.

Questionnaire pretest. We submitted the mailing cover letter, our directions for completing the questionnaire, and scale items to a pretest with four senior purchasing managers. Few items were slightly modified. The final list of value measures appears in the Appendix.

Data Collection

Sampling procedure. We collected data in a nationwide survey among purchasing managers of U.S. manufacturing firms. We used a three-wave mailing approach (Dillman 1978). We randomly selected 1950 members of ISM from the association's database. For an exploration sample, we contacted 527 managers, and for a validation sample, we contacted 1423 managers. We selected only senior-level managers, indicated by job titles such as Vice President Procurement, Director of Global Sourcing, Director of Supply Chain Management, or Purchasing Manager. In addition, we selected only manufacturing companies (Standard Industrial Classification codes 28–30 and 32–38).

A total of 118 (response rate of 22.4%) questionnaires were returned for the exploration sample, and 303 (response rate of 21.3%) were returned for the validation sample. We dropped 18 respondents who reported an average level of confidence in their responses of four or lower on a seven-point scale. We discarded 3 additional responses because of excessive missing data, which left a final sample of 112 for the exploration sample and 288 for the validation sample.

Sample characteristics. Respondents purchased a broad variety of components for multiple applications. Firms ranged from small enterprises to multibillion dollar companies. On average, customers had been buying from their main supplier for 13 years, with a standard deviation of 9 years. In our sample, on average, main suppliers captured 73.3% of customers' order volumes in a given product category. In contrast, second suppliers secured 19.5% of customers' needs. Finally, all other suppliers accounted for only 7.2% of orders.

Respondent characteristics. Respondents held senior positions in their firms. They averaged 17 years of experience in their area and 10.4 years with their companies.[1] Responses regarding confidence about answering the survey and knowledge of the supplier relationship displayed mean ratings of 6.0 (confidence in answering the survey), 6.01 (involvement in the supplier relationship), 6.15 (knowledge about the supplier), and 5.89 (influence of purchase decisions) on a seven-point scale.

Nonresponse bias. We assessed nonresponse bias following Mentzer, Flint, and Hult's (2001) recommended guidelines. We contacted a random sample of 30 nonrespondents over the telephone and asked them to answer our four questions that captured overall value perceptions in a supplier relationship (RELVAL 1–RELVAL 4 in the Appendix). In addition, we asked nonrespondents to provide background information on themselves and their company. The t-tests of group means revealed no significant difference between nonrespondents and our sample. Thus, nonresponse bias was not considered a problem in the present study.

Measurement Model

We report the results of our partial least squares (PLS) analysis in Figure 1 (for a detailed discussion of the PLS algorithm, see Lohmöller 1989). We chose PLS because of the formative nature of the higher-order value construct. This estimation procedure accommodates both reflective and formative measures (Fornell and Bookstein 1982). We measured relationship value on the basis of (1) the formative value dimensions identified in the exploratory study and (2) four reflective items that captured relationship value as a unidimensional concept on a high level of abstraction.

The formative and the reflective measurement approaches share 73% of their variance, indicating predictive validity of the formative value dimensions. All parameter estimates are significant at the 5% level. We cannot report an overall goodness-of-fit, because the objective of PLS is prediction versus fit (Fornell and Cha 1994). Partial least squares matches our research objective, that is, to assess the potential for differentiation of the formative value dimensions. Together with content validity established in the qualitative study by expert agreement, these results provide empirical evidence for construct validity (Rossiter 2002).

We measured relationship costs on three formative dimensions: direct product costs, acquisition costs, and operations costs. In line with Cannon and Homburg (2001), we summated the respective cost items to obtain a score for each of these cost drivers.

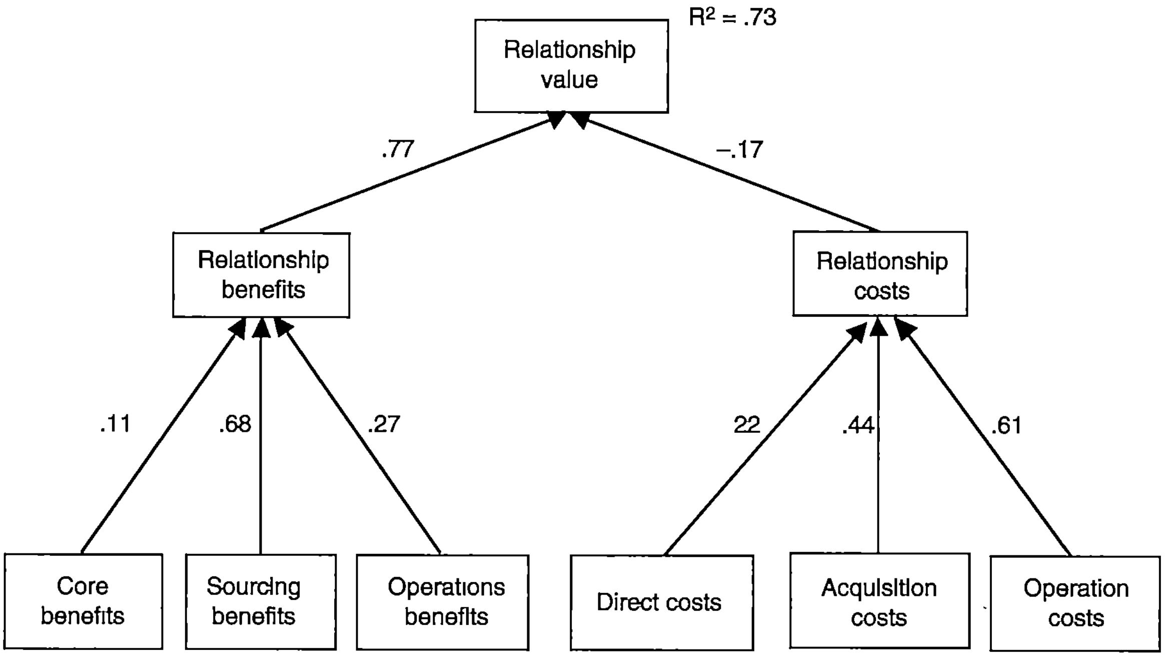
To operationalize relationship benefits, we developed a measurement model first. On the basis of our conceptual framework, we expected three dimensions to mirror Cannon and Homburg's (2001) conceptualization of relationship costs. To probe this conceptualization, we subjected the exploration sample to exploratory factor analysis. Three factors emerged on the basis of the Kaiser criterion (see Table 3). Product quality and delivery items formed the core benefits (24.4% variance explained). Personal and service benefit items loaded on a common factor that represented the benefits related to the sourcing process (28.4% variance explained). Know-how and time-to-market items loaded on a common factor that formed the operations benefits of a purchasing relationship (23.2% variance explained).

On the basis of the validation sample, we assessed convergent and discriminant validity (Table 4). Factor loadings, t-values, average variance extracted (AVE), and Cronbach's alpha are indicative of a high level of convergent validity. In addition, we assessed discriminant validity with Fornell and

---

[1] None of the background factors included in this research (i.e., company size, relationship length, and respondents' personal background and tenure with the company) moderated the results.

FIGURE 1
Formative Higher-Order Measurement Model



Notes: All parameter estimates are significant at the 5% level.

Larcker's (1981) criterion. Table 4 shows that the smallest AVE exceeds the squared correlation between each pair of the relationship value dimensions. This indicates a satisfactory level of discriminant validity. After we successfully established convergent and discriminant validity, we constructed summated scales for the benefit dimensions. Together with the relationship cost scores, they served as formative indicators for the first-order relationship benefit and cost dimensions, respectively.

Discussion and Implications

Against the background of supply base consolidation, our research contributes to a better understanding of how vendors of routinely purchased products earn the coveted main supplier position and defend their status against competitors. Overall, our study's findings suggest that a key supplier status indeed confers many substantial advantages to vendors. For example, key suppliers typically capture a significantly larger share of a customer's business than do other suppliers. In our study, on average, main suppliers secured 73.3% of customers' order volumes. In turn, second suppliers captured only 19.5% of customers' needs. Our study further revealed that a vendor's key supplier status may protect the firm from competition, at least in the short run. For example, customers will work with key suppliers to solve quality problems through joint teams. Similarly, main suppliers are granted an opportunity to adjust prices over time to match competition. Finally, key suppliers find themselves in a better position to compete for new business because they are typically involved early on in new product development.

While the advantages of working with fewer suppliers are well documented in the purchasing and supply chain management literature, scholars also recognize its downsides (Kekre, Murthi, and Srinivasan 1995; Monczka, Trent, and Handfield 2005; Newman 1989; Trevelen and Schweikhart 1988). Above all, buyers face increased dependence and, consequently, higher risks of supply disruption and supplier opportunism (Trevelen and Schweikhart 1988). Although suppliers gain greater volumes, improved consistency, and lower transaction costs, the cost of doing business with the customer may also increase (Newman 1989). Key suppliers are often required to provide internal cost data. Furthermore, they are permanently pressured to invest in the relationship. Finally, the main supplier position places a greater burden on the vendor to coordinate its own activities and subsuppliers (Monczka, Trent, and Handfield 2005, p. 278). Against this background, vendors must be aware that "gaining a greater share of a customer's business does a supplier very little good when that incremental business comes at the cost of reduced profitability" (Anderson and Narus 2003, p. 48).

Our research raised and answered several conceptual, measurement, and managerial issues. First, from a conceptual perspective, a thorough investigation of how value is created in a supplier relationship was necessary. Building on Cannon and Homburg's (2001) classification of relationship costs, we suggested an integrated framework that distinguishes between two fundamental value-creating dimensions and three levels at which these drivers operate. Following a grounded theory approach, we asked managers to identify and describe only those value drivers that allow a supplier to set itself apart from competition and gain key

TABLE 3
Exploratory Factor Analysis

Indicator	Factor 1 (Sourcing Process)	Factor 2 (Core Offering)	Factor 3 (Customer Operations)
Product1		.84	
Product2		.83	
Product3		.80	
Product4		.86	
Product5		85	
Product6		.85	
Service1	.66		
Service2	.72		
Service3	.70		
Service4	.77		
Delivery1		.72	
Delivery2		.73	
Delivery3		.75	
Know-how1			.64
Know-how2			.83
Know-how3			.70
Know-how4			.80
Know-how5			.87
Time to market1			.74
Time to market3			.75
Time to market4			.80
Personal1	.81		
Personal2	.84		
Personal3	.83		
Personal4	.80		
Personal5	.83		
Personal6	.73		
Personal7	.75		
Variance extracted (%)	28.4	24.4	23.2

supplier status. The results of our qualitative study suggest nine key differentiators in buyer–seller relationships: product quality, delivery performance, and direct product costs at the level of a supplier's core offering; service support, personal interaction, and acquisition costs in the sourcing process; and supplier know-how, time to market, and operations costs at the level of the customer's operations. To the best of our knowledge, this is the first conceptualization of relationship value grounded in managerial practice.

Second, from a measurement perspective, we developed items for each of the differentiators and consolidated them into an overall index of relationship value. The results of our cross-sectional study indicate that value creation in buyer–seller relationships can be not only conceptually described but also empirically measured. This is an important finding because scholars anticipated substantial difficulties in the empirical assessment of relationship value (Möller and Törrönen 2003; Parasuraman 1997; Wilson and Jantrania 1995). Because we cannot manage what we cannot measure, our research represents an essential step toward adding managerial relevance to this "problematic concept which cannot be ignored" (Wilson and Jantrania 1995, p. 63).

We measured relationship value in two complementary ways: (1) as a multidimensional construct and (2) as a unidimensional construct. To account for its conceptual breadth, empirical research should rely on formative multidimensional scales of relationship value. Reflective models of relationship value misconstrue the causal priority between the construct and its dimensions. However, they are appropriate if relationship value is considered a unidimensional construct on a high level of abstraction. Both measurement approaches share 73% of their variance. Consequently, researchers may choose between a multidimensional scale with multiple items and its unidimensional counterpart with only few items, depending on their research objectives. Relying on the multidimensional scale is key for research investigating the value-creating dimensions of a business relationship. For efficiency reasons, however, scholars might prefer the unidimensional scale when integrating value into a broader nomological network.

Third, from a managerial perspective, our research provides guidelines regarding each value driver's potential for differentiating among suppliers in a business relationship. To some extent, each value driver contributes to the explanation of observed variance in overall relationship value, and none of the differentiators should be neglected. However, Table 5 shows that the potential for differentiation in a buyer–seller relationship varies considerably among the various value drivers.

Our results suggest that relationship benefits display a stronger potential for differentiation than do cost considerations. Whereas relationship costs account for little more than 20% of the variance, relationship benefits explain nearly four times as much. This is a noteworthy finding. Conventional wisdom posits that customers emphasize cost considerations in business markets. Traditionally, the purchasing literature devotes particular attention to cost factors. For example, Anderson, Thomson, and Wynstra (2000, p. 310) argue that "[p]urchasing managers are more knowledgeable about using price and price changes as a basis for selecting product offerings than value and value changes." In addition, they recommend (p. 325) that "whenever possible, suppliers should make resource allocation trade-offs in favor of internal cost reductions, which allow the supplier to offer lower prices (or to postpone price increases), rather than performance enhancements."

Our findings provide a different picture, illustrating that vendor selection criteria are highly context dependent. Exploring value creation in ongoing relationships and conducting value assessments of physical products follow two different, yet complementary, logics. In the current study, we focused on key supplier relationships as units of analysis. From this perspective, relationship benefits and costs take on two very different roles. Whereas cost factors serve as key criteria to get a supplier on the short list of those vendors considered for business, relationship benefits dominate when deciding which supplier to name first among a set of available suppliers. Following this line of reasoning, cost competitiveness emerges as a necessary but not sufficient condition to gain key supplier status. In turn, offering superior benefits to the customer is essential for winning a substantial share of a customer's business.

In our examination of individual value drivers, we identified service support and personal interaction as core dif-

TABLE 4
Convergent and Discriminant Validity

A: Convergent Validity

Factor	Indicator	Loading	t-Value	AVE	α
Core offering	Product1	.83	32.75	75.0%	.96
	Product2	.91	70.72		
	Product3	.86	44.60		
	Product4	.86	33.03		
	Product5	.90	63.07		
	Product6	.91	63.79		
	Delivery1	.79	24.77		
	Delivery2	.86	49.61		
	Delivery3	.87	51.78		
Sourcing process	Service1	.83	32.26	73.8%	.96
	Service2	.83	30.38		
	Service3	.84	35.93		
	Service4	.85	38.61		
	Personal1	.88	46.76		
	Personal2	.90	69.03		
	Personal3	.88	48.51		
	Personal4	.90	53.32		
	Personal5	.88	52.64		
	Personal6	.84	38.32		
	Personal7	.82	26.55		
Customer operations	Know-how1	.75	23.53	72.6%	.97
	Know-how2	.86	37.08		
	Know-how3	.84	47.61		
	Know-how4	.87	56.21		
	Know-how5	.88	41.05		
	Time to market1	.86	44.91		
	Time to market3	.87	51.38		
	Time to market4	.87	53.71		

B: Discriminant Validity

	Core Offering	Sourcing Process	Customer Operations
Core offering	**.75**		
Sourcing process	.55	**.74**	
Customer operations	.38	.63	**.73**

Notes: Bold numbers on the diagonal show the AVE. Numbers below the diagonal represent the squared correlations.

TABLE 5
Value Drivers' Potential for Differentiation in Key Supplier Relationships

	Relationship Value Dimensions	
Sources of Value Creation	**Benefits**	**Costs**
Core offering	Core benefits: .11 × .77 = **.08**	Direct costs: .22 × .17 = **.03**
Sourcing process	Sourcing benefits: .68 × .77 = **.52**	Acquisition costs: .44 × .17 = **.07**
Customer operations	Operations benefits: .27 × .77 = **.21**	Operations costs: .61 × .17 = **.10**

Notes: Bold numbers show the variance explained by the formative value dimensions. We derive them by multiplying the path coefficients of the second-order dimensions by the coefficients of the respective first-order dimensions in Figure 1. Overall, benefits account for 8% + 52% + 21% = 81% of explained variance. In turn, costs account for 3% + 7% + 10% = 20% of explained variance. Benefits and costs do not add up to 100% because of rounding errors ferentiators. In the current study, these sourcing benefits offer the strongest potential for differentiation, explaining more than half of the observed variance (52%). They are followed by operation benefits. A supplier's know-how and its ability to improve a customer's time to market explain 21% of the observed variance. Operation costs (10% of observed variance), core benefits (i.e., product quality and delivery performance; 8% of observed variance), and acquisition costs (7% of observed variance) display a moderate potential for helping a firm gain and maintain key supplier status. Finally, direct product costs (i.e., price) show the weakest potential for differentiation in a key supplier relationship (3% of observed variance). These findings suggest that the core product and its price become less important differentiators in customer–supplier relationships. Offering value through personal interaction and service, access to know-how, and increased time to market has become important in securing a key supplier position.

However, note that managerial practice may well differ from our findings at the level of individual firms. For example, Noordewier, John, and Nevin (1990) suggest that repetitively used items and capital goods greatly differ in terms of acquisition and possession costs.[2] Indeed, repetitively used items involve significant order-handling and inventory management costs. In such a context, suppliers create value for customers by taking cost out of the sourcing process. In turn, capital goods require a thorough understanding of how the equipment helps customers reduce operation costs throughout the product's life cycle. In such an environment (i.e., in which the interpurchase intervals are much longer), managers base their decisions on payback periods and discounted cash flow analyses (Nordewier, John, and Nevin 1990). Nevertheless, our results represent a point of departure for managerial decision making. Because of the cross-sectional nature of the quantitative study, our findings provide a best-practice profile for companies striving for key supplier status. Our sample consists of winners only; that is, our analysis is based on data gathered from suppliers that successfully made their way to the top of the supply base. Consequently, our results deliver an informative benchmark for vendors of routinely purchased products that want to gain or maintain key supplier status.

On a final note, our study also provides interesting insights for secondary suppliers. By adopting a "Trojan horse" approach, a secondary supplier may break down an existing market offering into individual components and focus on a single product or service (Rust, Lemon, and Narayandas 2005, p. 313). Having established a foothold in a customer account, the supplier then can progressively expands its share. A secondary supplier may also develop a unique expertise and provide selected customers with a better solution at a lower cost. The supplier then can offer this capability for a reduced fee in exchange for an incremental share of a customer's business (Anderson and Narus 2003). Finally, a backup supplier can take advantage of the growing trend toward cross-sourcing among buying organizations (Newman 1989). Cross-sourcing not only helps customers mitigate supply risk but also opens a window of opportunity for backup suppliers to provide customers with the same product without needing to endure the burdens placed on the shoulders of the main supplier.

Limitations and Future Research Directions

As in any research project, the choices we made in this study imply limitations in the interpretation of our results. By focusing on understanding how a firm can differentiate itself from competition, we limited our assessment to a specific unit of analysis. Rather than showing how value is created in buyer–seller relationships overall, our findings refer to a particular context and research question.

First, we asked respondents to compare their main supplier with a second supplier for routinely purchased products. Such a focus excludes the examination of other supplier relations (e.g., value creation by "me-too" suppliers, sole supplier relationships). Similarly, our research did not investigate the sourcing of other product categories, such as capital goods.

Second, data for our research are based on the customer's perspective of value creation in key supplier relationships. A vendor's point of view might be different. Additional research could examine potential gaps between both parties' value perceptions.

Third, our research relied only on single respondents. Other members of the buying organization might emphasize different value drivers in a key supplier relationship. Because we cannot exclude potential biases, further research could address this limitation usefully.

Fourth, the choices we made in developing our sample frame warrant some caution in the generalization of our findings. We sampled respondents among members of the ISM. This approach might favor the inclusion of better-educated purchasing professionals and larger buying organizations.

Fifth, this research focused on dyadic business exchange relationships. However, buyer–seller relationships are embedded in complex networks of relationships (Achrol 1997; Anderson, Håkansson, and Johanson 1994; Ford 1990). Thus, further research could examine value creation in the broader network surrounding a customer and its key supplier.

Sixth, although we paid careful attention in the assessment of content and predictive validity, we did not raise the issue of nomological validity. Integrating value into the nomological network of relationship marketing remains a promising opportunity for further research (Hogan and Armstrong 2002, p. 19).

Finally, scholars have called for more longitudinal research on buyer–seller relationships (Narayandas and Rangan 2004). More must be learned about how value is created over time in close supplier–customer relationships. Further research using a longitudinal design could address this issue.

These limitations should be kept in mind when considering our results and their implications. Nevertheless, our findings provide new insights for academics and practitioners alike.

---

[2] We thank an anonymous reviewer for this insight.

APPENDIX
Operational Measures

		M	SD
Product Support			
Product1	Compared to the second supplier, the main supplier provides us with better product quality.	4.55	1.59
Product2	Compared to the second supplier, the main supplier meets our quality standards better.	4.62	1.60
Product3	Compared to the second supplier, the main supplier's products are more reliable.	4.47	1.59
Product4	Compared to the second supplier, we reject less products from the main supplier.	4.60	1.69
Product5	Compared to the second supplier, the main supplier provides us with more consistent product quality over time.	4.69	1.60
Product6	Compared to the second supplier, we have less variations in product quality with the main supplier.	4.53	1.68
Service Support			
Service1	Compared to the second supplier, the main supplier provides us with better services.	4.92	1.52
Service2	Compared to the second supplier, the main supplier is more available when we need information.	4.63	1.54
Service3	Compared to the second supplier, the main supplier provides us with more appropriate information.	4.83	1.65
Service4	Compared to the second supplier, the main supplier responds faster when we need information.	4.74	1.50
Delivery			
Delivery1	Compared to the second supplier, the main supplier performs better in meeting delivery due dates.	4.83	1.65
Delivery2	Compared to the second supplier, we have less delivery errors with the main supplier.	4.74	1.68
Delivery3	Compared to the second supplier, deliveries from the main supplier are more accurate (no missing or wrong parts).	4.77	1.69
Supplier Know-How			
Know-how1	Compared to the second supplier, the main supplier provides us a better access to his know-how.	4.68	1.59
Know-how2	Compared to the second supplier, the main supplier knows better how to improve our existing products.	4.73	1.55
Know-how3	Compared to the second supplier, the main supplier performs better at presenting us with new products.	4.81	1.52
Know-how4	Compared to the second supplier, the main supplier knows better how to help us drive innovation in our products.	4.77	1.69
Know-how5	Compared to the second supplier, the main supplier knows better how to assist us in new product development.	4.37	1.48
Time to Market			
Time to market1	Compared to the second supplier, the main supplier performs better in helping us improve our time to market	4.48	1.50
Time to market2	Compared to the second supplier, the main supplier helps us more in improving our cycle time.	4.55	1.57
Time to market3	Compared to the second supplier, the main supplier helps us more in getting our products to market faster.	4.40	1.58
Time to market4	Compared to the second supplier, the main supplier performs better in helping us speed up product development.	4.61	1.55
Personal Interaction			
Personal1	Compared to the second supplier, it is easier to work with the main supplier.	4.89	1.61
Personal2	Compared to the second supplier, we have a better working relationship with the main supplier.	5.07	1.52
Personal3	Compared to the second supplier, there is a better interaction between the main supplier's people and ours.	5.07	1.54
Personal4	Compared to the second supplier, we interact better with the main supplier.	4.91	1.57
Personal5	Compared to the second supplier, we can address problems more easily with the main supplier.	4.83	1.56
Personal6	Compared to the second supplier, we can discuss problems more freely with the main supplier.	4.69	1.59
Personal7	Compared to the second supplier, the main supplier gives us a greater feeling of being treated as an important customer.	4.72	1.63

APPENDIX
Continued

Relationship Value

		M	SD
RELVAL1	Compared to the second supplier, the main supplier adds more value to the relationship overall.	4.99	1.51
RELVAL2	Compared to the second supplier, we gain more in our relationship with the main supplier.	4.93	1.45
RELVAL3	Compared to the second supplier, the relationship with the main supplier is more valuable.	5.01	1.51
RELVAL4	Compared to the second supplier, the main supplier creates more value for us when comparing all costs and benefits in the relationship.	5.01	1.49

Relationship Costs

How do each of the following costs of the main supplier compare with the costs of your second-best supplier?

	Main Supplier's Costs Are Much Lower	Main Supplier's Costs Are Somewhat Lower	Main Supplier's Costs Are the Same	Main Supplier's Costs Are Somewhat Higher	Main Supplier's Costs Are Much Higher	M	SD
Purchasing price						2.51	1.02
Ordering costs						2.61	.72
Delivery costs						2.63	.89
Inventory carrying costs						2.60	1.81
Coordination and communication costs						2.52	.83
Manufacturing costs						2.62	.84
Downtime costs						2.57	.86

REFERENCES

Achrol, Ravi S. (1997), "Changes in the Theory of Interorganizational Relations in Marketing: Towards a Network Paradigm," *Journal of the Academy of Marketing Science*, 25 (1), 56–71.

Anderson, James C. (1995), "Relationships in Business Markets: Exchange Episodes, Value Creation, and Their Empirical Assessment," *Journal of the Academy of Marketing Science*, 23 (4), 346–50.

——— and David W. Gerbing (1991), "Predicting the Performance of Measures in a Confirmatory Factor Analysis with a Pretest Assessment of Their Substantive Validities," *Journal of Applied Psychology*, 76 (5), 732–40.

———, Håkan Håkansson, and Jan Johanson (1994), "Dyadic Business Relationships Within a Business Network Context," *Journal of Marketing*, 58 (October), 1–15.

———, Dipak C. Jain, and Pradeep K. Chintagunta (1993), "Customer Value Assessment in Business Markets: A State-of-Practice Study," *Journal of Business-to-Business Marketing*, 1 (1), 3–29.

——— and James A. Narus (1998), "Business Marketing: Understanding What Customers Value," *Harvard Business Review*, 76 (November–December), 53–65.

——— and ——— (2003), "Selectively Pursuing More of Your Customer's Business," *MIT Sloan Management Review*, 44 (Spring), 42–49.

——— and ——— (2004), *Business Market Management: Understanding, Creating, and Delivering Value*, 2d ed. Upper Saddle River, NJ: Pearson/Prentice Hall.

———, James B.L. Thomson, and Finn Wynstra (2000), "Combining Value and Price to Make Purchase Decisions in Business Markets," *International Journal of Research in Marketing*, 17 (4), 307–329.

Bendapudi, Neeli and Robert P. Leone (2002), "Managing Buiness-to-Business Customer Relationships Following Key Contact Employee Turnover in a Vendor Firm," *Journal of Marketing*, 66 (April), 83–101.

Bollen, Kenneth and Richard Lennox (1991), "Conventional Wisdom on Measurement: A Structural Equation Perspective," *Psychological Bulletin*, 110 (2), 305–314.

Cannon, Joseph P. and Christian Homburg (2001), "Buyer–Supplier Relationships and Customer Firm Costs," *Journal of Marketing*, 65 (January), 29–43.

Churchill, Gilbert A., Jr. (1979), "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research*, 16 (February), 64–73.

DeVellis, Robert F. (2003), *Scale Development: Theory and Applications*, 2d ed. Thousand Oaks, CA: Sage Publications

Diamantopoulos, Adamantios and Heidi M. Winklhofer (2001), "Index Construction with Formative Indicators: An Alternative to Scale Development," *Journal of Marketing Research*, 38 (May), 269–77.

Dillman, Don A. (1978), *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley & Sons.

Dwyer, Robert F. and John F. Tanner Jr. (1999), *Business Marketing: Connecting Strategy, Relationships and Learning*. Boston: Irwin/McGraw-Hill.

Flint, Daniel J., Robert B Woodruff, and Sarah Fisher Gardial (2002), "Exploring the Phenomenon of Customers' Desired Value Changes in a Business-to-Business Context," *Journal of Marketing*, 66 (October), 102–117.

Ford, David (1990), *Understanding Business Markets: Interaction, Relationships and Networks*. San Diego: Academic Press.

Fornell, Claes and Fred L. Bookstein (1982), "A Comparative Analysis of Two Structural Equation Models: LISREL and PLS Applied to Market Data," in *A Second Generation of Multivariate Analysis*, Claes Fornell, ed. New York: Praeger Publishers, 289–323.

——— and Jaesung Cha (1994), "Partial Least Squares," in *Advanced Methods of Marketing Research*, Richard P. Bagozzi, ed. Cambridge, MA: Blackwell, 52–78.

——— and David Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39–50.

Ganesan, Shankar (1994), "Determinants of Long-Term Orientation in Buyer–Seller Relationships," *Journal of Marketing*, 58 (April), 1–19.

Gassenheimer, Jule B., Franklin S. Houston, and J. Charlene Davis (1998), "The Role of Economic Value, Social Value, and Perceptions of Fairness in Interorganizational Relationship Retention Decisions," *Journal of the Academy of Marketing Science*, 26 (4), 322–37.

Glaser, Barney G. and Anselm L. Strauss (1967), *The Discovery of Grounded Theory*. Chicago: Aldine.

Hewitt, Kelly R., Bruce Money, and Subhash Sharma (2002), "An Exploration of the Moderating Role of Buyer Corporate Culture in Industrial Buyer-Seller Relationships," *Journal of the Academy of Marketing Science*, 30 (3), 229–39.

Hogan, John E. and Gary Armstrong (2002), "Toward a Resource-Based Theory of Business Exchange Relationships: The Role of Relational Asset Value," *Journal of Business-to-Business Marketing*, 8 (4), 3–28.

Jackson, Barbara B. (1985), "Build Customer Relationships That Last," *Harvard Business Review*, 63 (November–December), 120–28.

Jap, Sandy D. (1999), "Pie-Expansion Efforts: Collaboration Processes in Buyer–Supplier Relationships," *Journal of Marketing Research*, 36 (November), 461–75.

Jarvis, Cheryl Burke, Scott B. Mackenzie, and Philip M. Podsakoff (2003), "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research," *Journal of Consumer Research*, 30 (September), 199–218.

Kalwani, Manohar U. and Narakesari Narayandas (1995), "Long-Term Manufacturer–Supplier Relationships. Do They Pay Off for Supplier Firms?" *Journal of Marketing*, 59 (January), 1–16.

Kekre, Sunder, B.P.S. Murthi, and Kannan Srinivasan (1995), "Operating Decisions, Supplier Availability, and Quality: An Empirical Study," *Journal of Operations Management*, 12 (3–4), 387–96.

Krapfel, Robert E., Deborah Salmond, and Robert Spekman (1991), "A Strategic Approach to Managing Buyer-Seller Relationships," *European Journal of Marketing*, 25 (9), 22–37.

Lapierre, Jozée (2000), "Customer-Perceived Value in Industrial Contexts," *Journal of Business and Industrial Marketing*, 15 (2–3), 122–40.

Lohmöller, Jan Bernd (1989), *Latent Variable Path Modeling with Partial Least Squares*. New York: Springer.

Lyons, Thomas F., A. Richard Krachenberg, and John W. Henke Jr. (1990), "Mixed Motive Marriages: What's Next for Buyer-Supplier Relations?" *MIT Sloan Management Review*, 31 (Spring), 29–36.

McCracken, Grant (1988), *Qualitative Research Methods Series: The Long Interview*, Vol. 13, Newbury Park, CA: Sage Publications.

Mentzer, John T., Daniel J. Flint, and G. Tomas M Hult (2001), "Logistics Service Quality as a Segment-Customized Process," *Journal of Marketing*, 65 (October), 82–104.

Miles, Matthew B. and A. Michael Huberman (1994), *Qualitative Data Analysis: An Expanded Sourcebook*, 2d ed. Thousand Oaks, CA: Sage Publications.

Möller, Kristian and Pekka Törronen (2003), "Business Suppliers Value Creation Potential: A Capability-Based Analysis," *Industrial Marketing Management*, 32 (2), 109–118.

Monczka, Robert, Robert Trent, and Robert Handfield (2005), *Purchasing and Supply Chain Management*, 3d ed. Mason, OH: Thomson South-Western.

Narayandas, Das and Kasturi Rangan (2004), "Building and Sustaining Buyer–Seller Relationships in Mature Industrial Markets," *Journal of Marketing*, 68 (July), 63–77.

Newman, Richard G. (1989), "Single Sourcing: Short-Term Savings Versus Long-Term Problems," *Journal of Purchasing and Materials Management*, 25 (Summer), 20–25.

Noordewier, Thomas G., George John, and John R. Nevin (1990), "Performance Outcomes of Purchasing Arrangements in Industrial Buyer–Vendor Relationships," *Journal of Marketing*, 54 (October), 80–93.

Parasuraman, A (1997), "Reflections on Gaining Competitive Advantage Through Customer Value," *Journal of the Academy of Marketing Science*, 25 (2), 154–61.

Payne, Adrian and Sue Holt (1999), "A Review of the 'Value' Literature and Implications for Relationship Marketing," *Australasian Marketing Journal*, 7 (1), 41–51.

Rangan, V. Kasturi and George T. Bowman (1992), "Beating the Commodity Magnet," *Industrial Marketing Management*, 21 (3), 215–24.

Rossiter, John R. (2002), "The C-OAR-SE Procedure for Scale Development in Marketing," *International Journal of Research in Marketing*, 19 (4), 305–335.

Rust, Roland T., Katherine N. Lemon, and Das Narayandas (2005), *Customer Equity Management*. Upper Saddle River, NJ· Pearson/Prentice Hall.

Spekman, Robert E. (1988), "Strategic Supplier Selection· Understanding Long-Term Buyer Relationships," *Business Horizons*, 31 (July–August), 75–81.

Slater, Stanley R. (1997), "Developing a Customer Value-Based Theory of Marketing," *Journal of the Academy of Marketing Science*, 25 (2), 162–67.

Strauss, Anselm L. and Juliet Corbin (1998), *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2d ed. Newbury Park, CA: Sage Publications.

Trevelen, Mark and Sharon Bergman Schweikhart (1988), "A Risk/Benefit Analysis of Sourcing Strategies: Single Vs. Multiple Sourcing," *Journal of Operations Management*, 7 (4), 93–114.

Vandenbosch, Mark and Niraj Dawar (2002), "Beyond Better Products: Capturing Value in Customer Interactions," *MIT Sloan Management Review*, 43 (4), 35–42.

Walter, Achim, Thilo A. Müller, Gabriele Helfert, and Thomas Ritter (2003), "Functions of Industrial Supplier Relationships and Their Impact on Relationship Quality," *Industrial Marketing Management*, 32 (2), 159–69.

Webster, Frederick E., Jr. (1984), *Industrial Marketing Strategy*, 2d ed. New York. John Wiley & Sons.

Wilson, David T. (1995), "An Integrated Model of Buyer-Seller Relationships," *Journal of the Academy of Marketing Science*, 23 (4), 335–45.

——— and Swati Jantrania (1995), "Understanding the Value of a Relationship," *Asia-Australia Marketing Journal*, 2 (1), 55–66.

Woodall, Tony (2003), "Conceptualising 'Value for the Customer': An Attributional, Structural and Dispositional Analysis," *Academy of Marketing Science Review*, 12, 1–42.

Woodruff, Robert W. (1997), "Customer Value: The Next Source for Competitive Advantage," *Journal of the Academy of Marketing Science*, 25 (2), 139–53.

Markus Christen, Ganesh Iyer, & David Soberman

Job Satisfaction, Job Performance, and Effort: A Reexamination Using Agency Theory

The objective of this article is to clarify ambiguities in the literature regarding the relationships among three key constructs of work relationships: effort, job performance, and job satisfaction. The relationship between job performance and job satisfaction is of central interest to research in organizational psychology. However, empirical research in that area finds that the link between these constructs is weak at best. A negative effect of effort on job satisfaction is consistent with agency theory, but there is limited empirical evidence to support this assumption. Moreover, some studies have found a positive effect of effort on job satisfaction. Using a model that incorporates the main constructs from agency theory and organizational psychology, the current study finds a negative, direct effect of effort and a positive, direct effect of job performance on job satisfaction. The authors show that conflicting findings in the literature are the result of inconsistency in both the measurement and the definition of constructs across studies that do not fully account for all the relationships between constructs. The current findings emphasize the need to distinguish clearly between factors that represent employees' inputs in a work relationship (i.e., effort) and those that represent their outputs (i.e., job performance). The article also demonstrates the importance of properly accounting or controlling for all key variables to eliminate biases that can arise in empirical research on work relationships.

There is an extensive body of research in organizational psychology that considers the role of job satisfaction in managing effective work relationships. Similarly, job satisfaction is a widely studied construct in marketing research on sales force (Brown and Peterson 1993), retail store managers (Lusch and Serpkenci 1990), and service workers (Boyt, Lusch, and Naylor 2001). This research examines the antecedents of job satisfaction and, in particular, the effects of job performance, effort, and the compensation structure. However, findings in this literature about the relationships between job satisfaction and these antecedents have been inconsistent and even controversial. For example, despite the finding that people derive intrinsic value from work, the relationship between job performance and job satisfaction has been found to be inconsistent and weak (Brown and Peterson 1993; Iaffaldano and Muchinsky 1985). Similarly, studies that examine the effect of effort on job satisfaction find that it has a positive effect (Brown and Peterson 1994). This second finding appears to contradict the logic of the equally large literature on agency relationships in economics and marketing, which is based on the assumption that effort is costly to an agent and therefore reduces the agent's utility (or job satisfaction).

The objective of this article is to develop a model of work relationships to investigate the relationship between job satisfaction and its key determinants, job performance and effort. The premise is that a complete understanding of job satisfaction and work relationships must be predicated on a theory of how effort affects job satisfaction and the way that effort affects the relationship between job performance and job satisfaction. To this end, we draw on both agency theory and organizational psychology. We consider a role of effort that is consistent with a principal–agent model (Holmstrom 1979): When compensation and other factors are controlled for, effort is a cost for an agent. We then embed this cost in a job satisfaction model (Porter and Lawler 1968) to clarify the relationships between job satisfaction and its key antecedents. A clear understanding of these relationships is important to design employment contracts that optimize firm performance, while providing satisfactory incentives and compensation for an employee.

When we consider the three constructs of job satisfaction, job performance, and effort jointly, we find a strong positive effect of job performance on job satisfaction, a result that has long been hypothesized in marketing and organizational psychology but has received weak empirical support (Iaffaldano and Muchinsky 1985). In addition, when we account for the moderating effect of job performance, we find that effort has a strong negative effect on job satisfaction, a result that supports the theoretical assumption of empirical studies of contracts, such as sales force compensation plans (e.g., Lal, Outland, and Staelin 1994). Both omitted variables and construct definition problems explain why some previous research has found a posi-

---

Markus Christen is Associate Professor of Marketing (e-mail: markus.christen@insead.edu), and David Soberman is Associate Professor of Marketing (e-mail: david.soberman@insead.edu), INSEAD. Ganesh Iyer is Associate Professor of Marketing, Haas School of Business, University of California, Berkeley (e-mail: giyer@haas.berkeley.edu). The authors contributed equally to this research, and their names are listed in alphabetical order. The authors are deeply grateful to Ray Serpkenci for providing the data used in this study. They also thank Erin Anderson, Andy Mitchell, Gary Russell, Rick Staelin, and the participants of the joint INSEAD/ESSEC/HEC seminar and the Marketing Science conference at the University of California, Los Angeles, for their comments.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

tive effect of effort on job satisfaction. The result of our model and analysis is consistent with the agency theoretic view of effort. At the same time, we recognize the manner by which other determinants of job satisfaction mediate the impact of effort on job satisfaction.

Another important factor that must be incorporated into a model of work relationships is compensation. Agency theory makes a distinction between fixed compensation (e.g., salary) and variable compensation that depends on output (e.g., bonuses based on individual or firm performance). The agency theoretic prediction is that fixed compensation should have a significant effect on job satisfaction but not on effort, regardless of an employee's risk preference. Our results support this prediction.

In addition to an examination of fixed salary, our data set enables us to examine a common compensation variable that has received little attention in the literature: corporate-wide profit-sharing plans tied to overall firm performance. This type of compensation is interesting because in a large organization, it is neither fixed nor affected solely by the individual employee's performance. We find that these profit-sharing plans have a significant effect on both job satisfaction and effort.

This article is relevant to researchers in marketing and management who are interested in the relationship between job satisfaction and job performance. The results counter previous findings that there is no relationship between the two factors. Job performance has a direct and positive effect on job satisfaction when we properly account for effort. Effort has a positive effect on job performance and, thus, a positive, indirect effect on job satisfaction.

Next, the article demonstrates the problems of omitted variables and unclear construct definitions in researching work relationships within organizations. Findings about the relationship among effort, job performance, and job satisfaction can be unclear or inconsistent with received theory when key constructs are omitted or lumped together. This suggests the need for empirical research to distinguish clearly between factors that are inputs (i.e., effort) and those that are outputs (i.e., job performance).

We next provide a summary of the relevant literature, beginning with a discussion of the relationship between job performance and job satisfaction in marketing and organizational psychology. We follow this summary with a discussion of the effect of effort as it is purported in agency theory. On the basis of the literature review, we then set forth several hypotheses. We describe our empirical model and discuss the data and estimation approach. After we present the estimation results, we conclude with a discussion of the implications of our findings and highlight several important issues for further research.

Literature Review

The Relationship Between Job Performance and Job Satisfaction

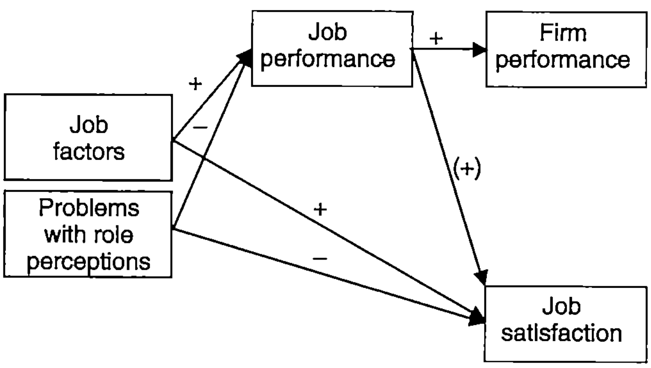
The idea that satisfied employees are more productive held through the 1970s. However, it was difficult to obtain support for the view that job satisfaction has a significant effect on job performance. As a result, the reverse (that an employee's job performance affects his or her job satisfaction) became the focus of research in the area (Lawler and Porter 1967). Although the idea that an employee's job performance affects his or her job satisfaction is consistent with several psychological theories, such as intrinsic motivation theory (Deci and Ryan 1985); few studies have found support for it (Iaffaldano and Muchinsky 1985). Similarly, organizational studies of the sales force in marketing invariably find that the relationship between job performance and job satisfaction is weak (Bagozzi 1980; Brown and Peterson 1993). As Brown and Peterson (1993) note, if the effect of job performance on job satisfaction is insignificant, firm actions designed to increase job performance should not have a direct effect on job satisfaction and related outcomes, such as employee turnover.

Conversely, there seems to be consensus that many antecedents of job satisfaction and job performance are common (e.g., effort, compensation, quality of supervision, clarity of job responsibilities). Consequently, it is essential to account for both direct and indirect effects of common antecedents to reach the conclusion that the effect of job performance on job satisfaction is insignificant.

The literature suggests that the antecedents of job satisfaction can be categorized into personal characteristics, role perceptions, and organizational variables (Brown and Peterson 1993). A typical job satisfaction model appears in Figure 1.

It is intuitive that effort and compensation are critical determinants for the viability of any employment situation. An employee will not be willing to exert effort unless he or she is paid, and the employer will not be willing to pay unless the employee works. A contract needs to be individually rational (i.e., both parties must expect to be better off by engaging in the work relationship). However, few, if any, studies include the employee's effort as an antecedent of job satisfaction (Brown and Peterson 1994). One exception is that of Clark and Oswald (1996), who use job satisfaction as a proxy for utility. They find a negative (albeit weak)

FIGURE 1
A Typical Job Satisfaction Model



Notes: Signs indicate the direction of effects according to the literature (Brown and Peterson 1993). Parentheses indicate a hypothesized relationship that has not been confirmed by empirical results.

effect of effort on job satisfaction. However, because job performance is omitted, its role as a potential mediator of the effect of effort on job satisfaction is not considered.

Many studies treat effort as part of job performance, which is defined broadly as an aggregate construct of effort, skill, and outcomes that are important to the employee and outcomes that are important to the firm (e.g., Behrman and Perreault 1984; Lusch and Serpkenci 1990; Walker, Churchill, and Ford 1977). Similar to the work of Bagozzi (1978), a few studies use a narrow definition of job performance based on actual sales or other objective productivity measures. However, these studies do not include effort as a separate construct.

We argue that it is important to define effort as distinct from job performance. From the perspective of an employee, job performance and effort are different. Effort is an input to work, and job performance is an output from this effort. From a firm's perspective, effort and job performance may be difficult to distinguish, and effort is often inferred from the output produced (the possibility of high effort and low output or low effort and high output is often not considered). This may explain the inclusion of effort in the definition of job performance. Some studies include work motivation as an antecedent to job satisfaction, but motivation ("I want to work hard") is not the same as exerted effort ("I did work hard and spent a lot of time and energy").

The implications of either neglecting effort or considering it a part of job performance for the empirically observed relationship between job performance and job satisfaction can be significant. If effort is costly for an employee, ignoring effort can bias the estimated effect of job performance on job satisfaction (because effort should increase job performance). Failing to control for effort induces a negative spurious correlation, which may reduce or even hide a true positive effect of job performance on job satisfaction. Similarly, by including effort in the measure of job performance, negative and positive aspects can nullify each other, yielding an effect for job performance that is again biased toward insignificance.

In summary, the existing job satisfaction research in marketing and organizational psychology is likely hampered by an omitted-variables bias and imprecise definitions of job performance. The objective of this article is to understand this relationship within a framework that overcomes these problems.

The Impact of Effort on Job Satisfaction

If effort is costly for an employee, it should have a negative, direct effect on job satisfaction. This implies that there is a conflict of interest between the employer, who wants the employee to work hard, and the employee, who wants the salary with the minimum possible effort. This conflict is the basis for the literature in economics on the relationship between principals and agents (Hart and Holmstrom 1987; Holmstrom 1979). Nevertheless, aside from the previously mentioned study by Clark and Oswald (1996), there is little direct empirical evidence that effort is a cost that makes employees less happy. Moreover, studies of job satisfaction that include effort as an antecedent tend to find that it has a positive effect on job satisfaction (Brown and Peterson 1994).

An objective of the article is to resolve the inconsistency between theoretical arguments and empirical findings. Our interest is to confirm empirically that the relationship between job satisfaction and effort is negative, because it is the basis for much of the research on contracts and job design. To accomplish this objective, we rely on prior research in economics that has established job satisfaction as a good proxy for utility (Clark and Oswald 1996; Friedman 1978) to untangle the direct and indirect effects of effort on job satisfaction (e.g., through compensation or job performance). The total effect of effort should indeed be positive, otherwise an employee would be better off not working.

Our conjecture is that the primary impediment to finding a negative relationship in previous empirical studies is a problem of omitted variables. Given that employment situations are typically subject to a problem of moral hazard, employers use a combination of outcome-based control (e.g., performance-based contracts) and behavior-based control (e.g., monitoring) to prevent employees from exerting minimal effort (Anderson and Oliver 1987), but these controls also affect employees' job satisfaction.[1] As a result, it is essential to account for the effect of these controls when estimating the valence and strength of the relationship between effort and job satisfaction. For example, in situations with incentive pay, which employers often use to compensate sales people, high effort (though costly) may lead to strong job performance and, thus, high compensation. This would increase job satisfaction and largely mitigate the negative effect of costly effort on job satisfaction. A model that does not account for the indirect path (through compensation) by which effort affects job satisfaction would find that the direct effect of effort on job satisfaction is either insignificant or even positive. A key requirement to identify the direct effect of effort on job satisfaction is to account for the indirect paths through which effort can affect job satisfaction (especially those created by the employer's control system). Next, we summarize the preceding discussion with a set of hypotheses and then propose a model of job satisfaction that incorporates an agency relationship and the key antecedents of job satisfaction into a comprehensive framework.

Hypotheses

The first hypothesis addresses the relationship between job performance and job satisfaction. By treating job performance as distinct from effort and accounting for the direct effect of effort on job satisfaction, we expect the following relationship:

> H_1: All else being equal, an employee's job satisfaction increases with his or her job performance.

---

[1] Moral hazard is a contracting problem between a principal and an agent when an agent expends effort that only he or she observes and the effect of effort on output is uncertain. Consequently, the agent will not choose an efficient level of effort.

In the context of our organizational setting, we show how null effects can be found by adopting approaches used in previous job satisfaction studies.[2]

H_2: All else being equal, omitting effort as a separate antecedent of job satisfaction, either by excluding it from the model or by combining it with job performance, reduces the effect of job performance on job satisfaction.

The next three hypotheses are based on the existence of a typical agency relationship between the firm and an employee.

H_3: All else being equal, job performance increases with an employee's effort and ability.

H_4: All else being equal, firm performance increases with an employee's job performance.

A basic assumption of agency theory is that effort is costly to an agent. By defining effort as an input and by considering indirect paths through which effort affects job satisfaction, we hypothesize that effort should be costly for an employee:

H_5: All else being equal, an employee's job satisfaction decreases with his or her effort.

Our last hypothesis posits a direct implication of the agency model for fixed compensation (i.e., compensation that does not change with an employee's output).

H_6: All else being equal, fixed compensation has a significant, positive effect on an employee's job satisfaction but no significant effect on the employee's effort.

Model

The firm's objective is to maximize financial performance within the business unit or department affected by the employee's job performance. Similarly, the employee maximizes job satisfaction. Business unit performance is assumed to be, in part, a result of the employee's job performance, which is determined by the employee's effort and ability. Our data consist of the performance and satisfaction of store managers in a retail chain, so the relevant business unit for each employee is a store. The firm (in our case, the retail chain) controls compensation and promotion opportunities and can influence other aspects of the job (e.g., responsibilities). The idea is that these variables potentially influence both the employee's job satisfaction and the employee's willingness to work hard. We summarize the model in Figure 2.

The model translates into a system of four equations that represent (1) the relationship among store performance, SP_i; the employee's job performance, JP_i; and the employee's effort, ME_i (Equations 1–3) and (2) the relationship between job satisfaction, JS_i, and relevant aspects of the employee's work situation, including effort and job performance (Equation 4):

$$(1) \qquad SP_i = \alpha_1 + \beta_{JP1}JP_i + \varepsilon_{1i},$$

$$(2) \qquad JP_i = \alpha_2 + \beta_{ME2}ME_i + \beta_{MA2}MA_i + \varepsilon_{2i},$$

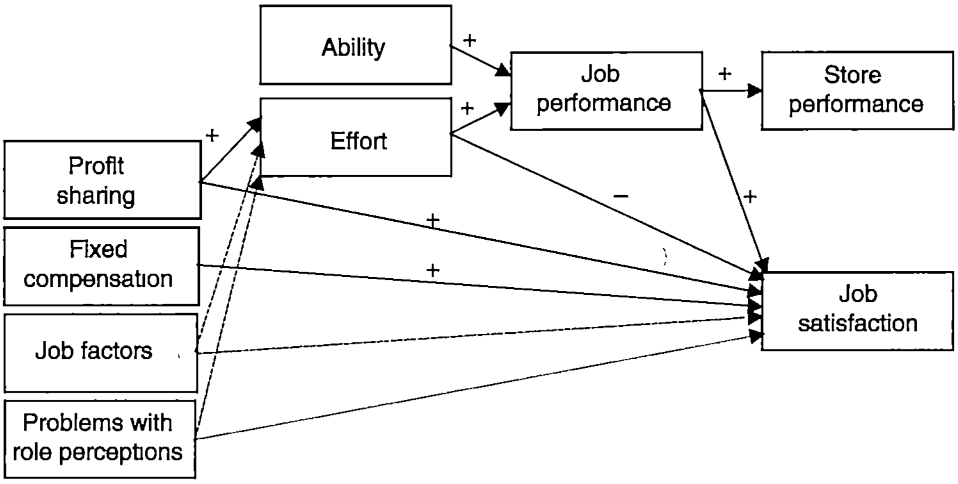
$$(3) \qquad ME_i = \alpha_3 + \beta_{C3}C_i + \beta_{PS3}PS_i + \gamma_{j3}X_{ji} + \delta_{k3}Y_{ki} + \varepsilon_{3i},$$

and

$$(4) \quad JS_i = \alpha_4 + \beta_{ME4}ME_i + \beta_{JP4}JP_i + \beta_{C4}C_i + \beta_{PS4}PS_i + \gamma_{j4}X_{ji} + \delta_{k4}Y_{ki} + \varepsilon_{4i},$$

where C_i is employee i's perception of compensation and PS_i is a dummy variable indicating participation in the

FIGURE 2
Proposed Model of Job Satisfaction and Store Performance



Notes: Signs indicate the direction of hypothesized effects. The dashed lines indicate relationships included in the model without specifying hypotheses profit-sharing plan. Following the literature, we divide the aspects of the job into job characteristics, X_{ji}, and role perceptions, Y_{ki}. Job characteristics include the employee's perception of the job's attractiveness, the job's autonomy, the quality of supervision, and the social climate at the work place. Role perceptions refer to the employee's perception of his or her responsibilities in terms of clarity, degree of overload, and degree of conflict. The nine β parameters are estimated to assess our hypotheses. We subsequently define and describe the four endogenous constructs and the key exogenous factors (we discuss measurement issues in the "Data and Estimation" section).

---

[2] The positive effect of job performance on job satisfaction appears at odds with standard agency theory. However, a positive effect may be consistent with a model of repeated interactions between a principal and an agent.

The store performance equation (Equation 1) reflects a vertical relationship in which the employee (in this case, the store manager) performs different tasks to generate outputs for the firm (the retail chain), and SP_i is a measure of a store's financial performance. The summary construct, JP_i, is a measure of the employee's job performance. We define this as a qualitative assessment of a store's performance relative to targets and objectives set for that store (as summarized by a supervisor). We treat job performance as an intermediate outcome variable (between effort and store performance), but our framework also recognizes that other factors affect store performance.[3]

The job performance equation (Equation 2) reflects the idea that the employee's job performance, JP_i, is a function of his or her effort, ME_i. Consistent with the work of Campbell and Pritchard (1976), we define effort as the amount of energy and time an employee puts into the job. Employees are not endowed with the same ability to perform the job, MA_i (ability refers to an employee's skill and knowledge related to the specific duties of the job). This should also have a significant effect on observed job performance.

The effort equation (Equation 3) reflects the assumption that effort, ME_i, is influenced by various job factors, some of which the firm controls. The set of job factors included in Equation 3 consist of compensation, C_i; access to the profit sharing plan, PS_i; four job characteristics, X_{ji}; and three role perception factors, Y_{ki} (we elaborate on these characteristics and factors subsequently). We used a dummy variable to indicate an employee's participation in the profit-sharing plan (29% of the store managers in the sample were included in the plan).

Finally, the job satisfaction equation (Equation 4) is based on the assumption that an employee's job satisfaction is a function of his or her job performance, his or her effort, and the job characteristics and role perceptions that affect his or her effort (see the discussion of Equation 3).

Our model includes four job factors, X_{ji} ($j = 1, \ldots, 4$), that are standard in the literature: job autonomy, job attractiveness, quality of supervisory feedback, and social climate. Together with the compensation variables, they represent the job's core characteristics (Fried and Ferris 1986). We define "job autonomy" as the degree to which the firm provides independence and discretion to the employee in fulfilling his or her role. We define "job attractiveness" as the degree to which a job is exciting, challenging, and provides a sense of accomplishment. We define "supervisory feedback" as the extent to which the employee receives information about his or her performance. Supervisory feedback is important because it often has psychological value for an employee. It is also a key element of the firm's control system. Finally, we define the variable "social climate" as the degree to which there is a good working relationship among the employees within the store. We use the employee's perception of these factors because perceptions are the primary drivers of the effort taken and satisfaction received (see, e.g., Judge, Bono, and Locke 2000).[4]

We include three role perception variables, Y_{ki} ($k = 1, \ldots, 3$): role ambiguity, role conflict, and role overload. These are known to have significant, negative effects on both job performance and job satisfaction (Brown and Peterson 1993). We do not have a direct interest in the effect of these variables, but we include them in the model because of their acknowledged importance in explaining job performance and job satisfaction. We define "role ambiguity" as a combination of uncertainty about the relationship between action and output and a lack of clear directions and behavioral requirements from the supervisor (Rizzo, House, and Lirtzman 1970). We define "role conflict" in terms of the congruency of various job responsibilities (i.e., the degree to which goals, objectives, and responsibilities of a position conflict with one another; Rizzo, House, and Lirtzman 1970). Role conflict can arise from inconsistent demands from different role partners (e.g., supervisors, coworkers, customers) on the employee; incongruencies with personal values; conflicts among different roles; and conflicts among time, resources, and demands (role overload). Because the workload of the employees in our data set is high and a distinctive job attribute, we include "role overload" as a separate factor.

To complete the model specification, we include intercepts, α_ℓ, and error terms, $\varepsilon_{\ell i}$, for each of the $\ell = 1, \ldots, 4$ equations to account for unobserved factors and random measurement error. We also add a set of variables that capture store and store manager characteristics for control and identification purposes. We provide details in the discussion of our estimation approach. Next, we describe the data set and discuss measurement and estimation issues.

Data and Estimation

Data Set

We use data that Lusch and Serpkenci (1990) collected for their study of the effect of personal difference variables on the job satisfaction of retail store managers. This data set is based on a survey of a U.S. grocery retailer with more than 200 supermarkets. It contains typical measures collected for job satisfaction studies and extensive operating and accounting information for each store. The stores are uniform in terms of merchandise and layout, and most stores have 7000 to 10,000 square feet of selling space. Stores are about equally distributed over metropolitan and rural areas. Each store is staffed by a store manager and one or more assistant managers. A district supervisor monitors the operations of 10 to 15 stores and reports directly to senior management.

---

[3] When the performance of a firm depends on various other exogenous factors, it is important to control for bad (good) outcomes, even when a manager's job performance is high (low). Therefore, eliminating job performance from the model and using only store performance could give a misleading result for the effect of effort. If the most talented and hardest working store managers were assigned to the worst-performing stores to improve them, it might be possible to find a negative effect of effort on store performance.

[4] Effort and job satisfaction may also be affected by promotion opportunities. A factor analysis of the data (see the "Data and Estimation" section) indicates that "promotion" cannot be identified independently of compensation. Thus, we exclude promotion as an independent factor from the analysis.

The data consist of three individual data sets collected from three different sources. The retailer provided operating and financial data for each store for the most recent three years. District managers (supervisors) completed a survey to provide an assessment of store managers' performance, effort, and ability. Store managers completed a survey on the quality of their work life, which provides the data for the store managers' assessment of job factors, role perceptions, and job satisfaction. This survey also captured several of the store managers' personal and attitudinal characteristics.

For the statistical analysis, there were 188 usable observations from 226 stores (18 surveys were not returned, and 20 were eliminated because of missing data). In addition, we eliminated 11 observations from store managers who were at the current job for less than one year, which left us with 177 observations for estimation.

Measurement of Constructs

Job satisfaction. Locke (1976, p. 1300) defines job satisfaction as "a pleasurable or positive emotional state," which is "a function of the perceived relationship between what one wants from a job and what one perceives it is offering." In essence, job satisfaction is an overall state that is derived from experiencing a work situation. Because our approach embeds a principal–agent structure within a job satisfaction model, we use job satisfaction as a proxy for the utility from working (derived by the employee). This assumption follows other "workplace studies" that have an economic orientation (Clark 1997; Clark and Oswald 1996; Friedman 1978).[5]

There are two general approaches to measure overall job satisfaction. The first determines job satisfaction by the measurement of satisfaction with different job "facets" (e.g., Smith, Kendall, and Hulin's [1969] job descriptive index); the second measures overall or "global" job satisfaction directly (e.g., Hackman and Oldham's [1975] job diagnostic survey). Research shows that in many cases, the global approach is superior to the facet approach (e.g., Scarpello and Campbell 1983). As a result, we use a global three-item measure of job satisfaction.

Store manager effort, ability, and job performance. We obtained measures of effort, ability, and job performance (for each store manager) from two different scales completed by the district supervisor. The Appendix provides a summary of the different scale items we used to measure the different constructs in our model. As we show in Table 1, all measures have good reliability properties. Cronbach's alpha values range from .81 for ability to .89 for effort.[6]

Store performance. The economic outcome of interest to the retail chain is store profits. To reduce potential accounting problems, we use a store's operating profit, which we define as store sales less store operational expenses for the current year. To control for store size, we divide the operating profit measure by available store selling space for the same year.[7]

Compensation, job factors, and role perceptions. The scales that Lusch and Serpkenci (1990) use to measure a store manager's perception of compensation and job factors are derived from the job descriptive index (Smith, Kendall, and Hulin 1969). They contain sufficient items to measure a store manager's perceptions of the five factors reliably (i.e., compensation, job attractiveness, job autonomy, supervisory feedback, and social climate). Individual store managers provided answers. Each of the five factors is measured by its respective scores on two to four items based on a five-point Likert scale, ranging from "strongly agree" to "strongly disagree," or on a six-point Likert scale, ranging from "definitely yes" to "definitely no." The scale items appear in the Appendix. Table 1 shows that Cronbach's alpha values range from .71 for job autonomy and social climate to .85 for compensation. As we previously noted, we indicated inclusion in the corporate profit-sharing plan with a dummy variable. Lusch and Serpkenci (1990) measure the three different role factors with a multi-item scale (see the Appendix). Cronbach's alpha for these three factors ranges between .69 for role conflict and .80 for role overload.

To ensure that the factors measure unique aspects of a store manager's job situation, we conducted a factor analysis. An exploratory factor analysis yielded eight factors with eigenvalues greater than one. We then conducted a confirmatory factor analysis to test the eight-factor solution, and we compared it with other possible solutions. Table 2 provides a summary of various goodness-of-fit indicators for different factor solutions. Although the chi-square test statistic indicates that the restrictions underlying the eight-factor solution are significant (which is expected given the

---

[5] These studies define a worker's utility function as u = u(y, e, i, j), where y is income, e is effort extended, and i and j are sets of individual and job parameters, respectively. In contrast to this specification, however, we also include job performance as a determinant of utility.

[6] Effort, ability, and job performance are significantly correlated (see Table 1). Nevertheless, a factor analysis of the data strongly indicated the presence of three factors. For each construct, we selected scale items on the basis of the results of the factor analysis. Moreover, higher correlation hurts efficiency, but estimates are still consistent.

[7] Note that unobserved store-specific factors and the retailer's strategy can affect both store performance and a store manager's effort. This can lead to biased parameter estimates if not properly controlled for. By measuring job performance relative to store objectives, we can account for heterogeneity in profit potential across stores. We also include several store characteristics in the estimation to mitigate this problem further.

TABLE 1
Pairwise Correlation Coefficients and Measurement Properties of Variables

	Variables													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Compensation														
2. Profit sharing[a]	.02													
3. Attractiveness	.35	−.04												
4. Autonomy	.38	.00	.42											
5. Supervisor	.37	.14	.37	.38										
6. Social climate	.48	.05	.53	.45	.36									
7. Role ambiguity	−.28	−.03	−.39	−.46	−.46	−.30								
8. Role conflict	−.11	.05	−.13	.06	−.11	−.12	.40							
9. Role overload	−.28	.04	−.37	−.24	−.30	−.24	.40	.35						
10. Ability	.13	.09	−.02	.13	.10	−.03	−.07	−.04	−.16					
11. Manager effort	.06	.10	−.02	.17	.15	−.06	−.04	.08	−.10	.69				
12. Job performance	.14	.02	−.07	.08	.11	−.03	−.16	−.05	−.17	.66	.66			
13. Job satisfaction	.47	.10	.58	.43	.27	.62	−.34	−.16	−.34	.10	.08	.09		
14. Store performance[b]	06	−.07	−.04	−.13	.01	−.08	−.04	−.10	−.19	.35	.19	.33	−.01	
M	1.88	.29	3.74	4.13	2.99	3.54	1.87	2.11	2.75	4.45	4.39	4.30	4.53	56.8
SD	.77	.46	.75	.85	.73	.87	.77	.88	.79	.89	.78	1.21	1.19	21.8
Number of scale items	4	1	4	2	4	2	4	3	3	3	5	3	3	—
Cronbach's α	.85	—	.75	.71[c]	.73	.72[c]	.79	.69	.80	81	.89	.88	.67	—

[a] Indicator variable
[b] Store operating profit relative to selling space.
[c] Simple correlation

TABLE 2
Goodness-of-Fit Statistics for Different Factor Solutions

Number of Factors[a]	Steiger–Lind RMSEA	Chi-Square (d.f.)	Jöreskog GFI	Bentler Comparative Fit Index	Parsimonious Fit Index	Bollen's Rho	Bollen's Delta
3	.116	998.7 (321)	.696	.695	.559	.575	.698
4 (Overload)	.101	918.7 (319)	.736	.730	.584	.607	.733
5 (Social climate)	.090	810.9 (315)	.765	.777	.614	.648	.780
6 (Autonomy)	.077	714.5 (310)	.793	.818	.638	.685	.820
7 (Supervisor)	.062	628.3 (304)	.825	.869	.654	.728	.866
8 (Conflict)	.049	481.6 (297)	.875	.912	.685	.764	.904
9 (Promotion)	.048	474.1 (289)	.883	.907	.690	.768	.899

[a] The three factors are compensation (compensation, promotion, social climate), work (attractiveness, role conflict, role overload), and organization (autonomy, role ambiguity, supervisor). The subsequent rows indicate higher-factor solutions. The factor that is separated from the previous solution with fewer factors is in parentheses.

Notes: RMSEA = root mean square error of approximation, and GFI = and goodness-of-fit index

sample size), its goodness-of-fit increases significantly compared with solutions with fewer factors. In addition, the indicators have reasonably high values.

In summary, we use standard measures for the constructs of interest, and they satisfy typical reliability standards. Because our data contain several five-, six-, and seven-point Likert scales, we standardized all item responses before creating indicators for estimation.

Estimation Approach

The model we outlined in the previous section is a triangular system of four simultaneous equations with four endogenous variables (effort, job performance, job satisfaction, and store performance), ten exogenous variables (a store manager's ability, six job factors, and three role perception variables), and several store and store manager characteristics added for control and identification purposes. To simplify the exposition, we do not discuss them in detail.[8] However, the error terms across the four equations are most likely correlated (i.e., the error covariance matrix is not diagonal, that is, the system is not fully recursive). In this situation, Lahiri and Schmidt (1978) show that a triangular system can be estimated in the manner of seemingly unrelated regression models (ignoring the simultaneity) with maximum likelihood estimation. This approach yields consistent and efficient estimates of the error variances, which is necessary for parameter estimates to be consistent (Greene 1991).[9]

---

[8] We included the following variables in our system of equations: store performance equation (Equation 1): age of the store, number of store employees, tenure of the store manager, and a dummy variable for a urban store location; job performance equation (Equation 2): a series of dummy variables to indicate the supervisor; effort equation (Equation 3): a dummy variable for urban store location, tenure, and age of store manager and four personal difference variables we adapted from Lusch and Serpkenci's (1990) study; and job satisfaction equation (Equation 4): four personal difference variables. Thus, for every equation, the number of exogenous variables is larger than the number of included endogenous variables, which satisfies the order condition of identification. The restriction that effort and job satisfaction do

One estimation problem is the possibility that access to the profit-sharing plan is determined through self-selection by the store managers or by the chain. A comparison of demographic and personality variables between store managers with and without the plan suggests that enrollment in the plan was random. A second problem is the possibility of random shocks that can simultaneously affect job performance and store performance. This can lead to a problem of contemporaneous correlation in the store performance equation (Equation 1), which we control for by using instrumental variable estimation. Specifically, we use store performance and other available variables lagged by one and two years as instruments. We then use the predicted value of store performance to estimate the system of equations.

Results

We first present the results for our four-equation model, which yield tests of H_1 and H_3–H_6. We then analyze two alternative models to replicate several controversial results in job satisfaction studies. These alternative models enable us to assess H_2.

Effort, Job Performance, and Compensation

The estimation results for Equations 1–4 appear in Table 3. Rows 2–4 show the results for the direct and indirect effects of a store manager's effort and job performance, Rows 5–6 show the results for the compensation variables, Rows 7–10 show the results for the effects of job characteristics, and the bottom three rows show the effect of role perceptions.

First, the effect of job performance on the manager's job satisfaction is positive and highly significant, providing strong support for H_1 ($\beta_{JP4} = .286, p < .001$).[10] The data can also be used to refute the idea that happy employees are more productive. If we reverse this relationship by adding job satisfaction to Equation 2 and by removing the effect of job performance from Equation 4, we find no significant effect of job satisfaction on job performance ($\beta = .098, p = .20$).

Second, as we proposed in H_3, job performance increases with both effort ($\beta_{ME2} = .158, p < .05$) and ability ($\beta_{MA2} = .387, p < .001$). Consistent with H_4, store performance increases with job performance ($\beta_{JP1} = 6.67, p < .001$). We also conducted a series of tests to assess the robustness of our findings with respect to alternative assumptions. For several variables, we tested for nonlinear effects on job satisfaction. All were insignificant. The positive effect of job performance on store performance is robust with respect to different store performance measures (e.g., net profit, sales per square foot). Because of unob-

---

not directly affect store performance ensures that the rank condition is satisfied as well.

[9] Three-stage least squares estimation can provide more efficient results by imposing greater constraints on the system. However, this makes estimates more susceptible to model specification problems. In any event, the results generated with three-stage least squares estimation lead to identical conclusions (the magnitudes of some estimated parameters are different).

[10] When our hypotheses make clear directional predictions, we report the statistical significance based on one-tailed t-tests. We report the significance of effects of other parameters based on two-tailed t-tests. Estimation details for additional analyses we presented in this section are available on request.

TABLE 3
Estimation Results

Parameter	Store Performance (Equation 1)	Store Manager Job Performance (Equation 2)	Store Manager Effort (Equation 3)	Store Manager Job Satisfaction (Equation 4)
Intercept (α_ℓ)	27.0*** (4.93)	–.495** (.211)	.078 (.213)	–.054 (.051)
Effort ($\beta_{ME\ell}$)		.158** (.074)		–.269*** (.061)
Ability ($\beta_{MA\ell}$)		.387*** (.073)		
Job performance ($\beta_{JP\ell}$)	6.67*** (1.33)			.286*** (.054)
Compensation ($\beta_{C\ell}$)			–.001 (.087)	.168*** (.057)
Profit sharing ($\beta_{PS\ell}$)			.183* (.109)	.246*** (.080)
Job autonomy ($\gamma_{1\ell}$)			.249*** (.079)	.096* (.053)
Job attractiveness ($\gamma_{2\ell}$)			.020 (.092)	.476*** (.061)
Supervisory feedback ($\gamma_{3\ell}$)			.140* (.080)	–.128** (.060)
Social climate ($\gamma_{4\ell}$)			–.204** (.083)	.176*** (.056)
Role ambiguity ($\delta_{1\ell}$)			.093 (.091)	–.026 (.060)
Role conflict ($\delta_{2\ell}$)			–.024 (.054)	–.047 (.031)
Role overload ($\delta_{3\ell}$)			–.119* (.070)	–.083* (.045)

*$p < .10$ (two-tailed t-test).
**$p < .05$ (two-tailed t-test).
***$p < .01$ (two-tailed t-test).
Notes. Number of observations = 177; total number of parameters estimated (including control variables) = 67. Standard errors are in parentheses.

served contemporaneous effects, we used an instrumented store performance variable in our estimations. The actual measure of store performance can artificially inflate the measured effect of job performance on store performance. We repeated the estimation with actual store performance and found this to be the case. To determine whether a manager's job satisfaction depends on the performance of his or her store (Ostroff 1992), we added store performance to Equation 4. We found that the effect of store performance on job satisfaction is insignificant ($\hat{\beta} = .002, p = .21$).

Third, and most important, we found strong support for H_5 (i.e., effort is a "cost" to the store manager; $\hat{\beta}_{ME4} = -.269, p < .001$). Finally, the estimate for the effect of compensation on job satisfaction has the expected positive sign and is statistically significant ($\hat{\beta}_{C4} = .168, p < .01$), whereas its effect on effort is not significant ($\hat{\beta}_{C3} = -.001, p = .99$), as we expected. These two results support H_6. We also estimated the system with an estimate of the store manager's annual salary added to Equations 3 and 4 and found that the results are unaffected. For job satisfaction, the effect of actual salary is only marginally significant.

As we previously mentioned, our data set provides us with the ability to examine the effects of a compensation component that is used frequently in firms but has not been subject to significant empirical research: corporate-level profit-sharing plans. As with other compensation elements, access to this plan, which some store managers had, should increase job satisfaction (all else being equal). As we show in Table 3, we find a positive effect of the profit-sharing plan on job satisfaction ($\hat{\beta}_{PS4} = .246, p < .001$). The effect of this plan on effort is less obvious. According to agency theory, there should be no effect, because the effort of an individual store manager has only a small impact on corporate profitability. Conversely, arguments in the organizational psychology literature suggest a positive effect because contingent rewards that depend less directly on a person's own performance are considered less controlling (Deci 1971; Ryan, Mims, and Koestner 1983). We found that access to the profit-sharing plan leads to greater effort ($\hat{\beta}_{PS3} = .183, p < .05$). Therefore, it would be useful to investigate further the mechanism through which such plans affect the agent's effort.

By carefully separating input (store manager's effort) from output (store manager's job performance), we find support for both the importance of intrinsic rewards that store managers obtain through job performance (H_1) and the assumption that effort per se is a disutility or cost for them (H_5). Because a store manager's effort increases job performance (H_3), it has a positive, indirect effect on job satisfaction. However, this positive, indirect effect of effort ($.158 \times .286 = .045$) is of smaller magnitude than the negative, direct effect (–.269). This implies that a store manager must receive compensation for his or her effort.

Other Job Factors and Role Perceptions

The estimates for the four job factors and the three role perceptions included in the job satisfaction equation (Equation 4) have signs that are consistent with existing research, except for the effect of the supervisor, which is negative. A store manager's job satisfaction increases with job autonomy ($\hat{\gamma}_{14} = .096, p < .1$), job attractiveness ($\hat{\gamma}_{24} = .476, p < .001$), and good social climate ($\hat{\gamma}_{44} = .176, p < .01$). In addition to the quality of supervisory feedback ($\hat{\gamma}_{34} = -.128, p < .05$), the three role perception factors also decrease job satisfaction, though at best, their effects are only marginally significant (ambiguity: $\hat{\delta}_{14} = -.026, p = .67$; conflict: $\hat{\delta}_{24} = -.047, p = .13$; overload: $\hat{\delta}_{34} = -.083, p < .1$).

For the store manager effort equation (Equation 3), we found (expected) positive effects of job autonomy ($\hat{\gamma}_{13} = .249, p < .01$) and the quality of supervisory feedback ($\hat{\gamma}_{33} = .140, p < .1$). Conversely, job attractiveness has no significant effect on effort ($\hat{\gamma}_{23} = .020, p = .83$). Notably, we found that a better social climate is associated with lower effort ($\hat{\gamma}_{43} = -.204, p < .05$). This finding indicates that in hierarchical organizations, such as a retail store, excessive socialization might interfere with the retailer's objectives of extracting high effort from the store employees. Although the store manager gains satisfaction from positive work relationships, having subordinates that are more than just employees (i.e., friends) may interfere with the ability to manage them. This counters the argument that retailers that create close working relationships between employees and store management have higher levels of performance (Dunne and Lusch 1999). Of the three role perception variables, only the effect of role overload is marginally significant ($\hat{\delta}_{34} = -.119, p < .1$).[11] In general, we find that the effect of the role perception variables are weaker than reported in other job satisfaction studies (Brown and Peterson 1993; Lusch and Serpkenci 1990). This difference may be caused by the exclusion of effort, as was the case in these studies (i.e., the disutility from effort could have been partially attributed to stress factors).

As we mentioned previously, we included several variables that were indicative of store and store manager characteristics in the estimation for control and identification purposes. We do not discuss them except for noting that the signs of the personal difference variables are consistent with Lusch and Serpkenci's (1990) hypotheses.

Alternative Models

We now examine how the results of our analysis are affected by changing the definition of job performance and excluding effort as an antecedent from the model (as H_2 suggests). To test these ideas, we eliminate effort (i.e., the effort equation) from the model and add the job characteristic and role perception variables to the job performance equation. In a first alternate model, we use an "aggregate" measure of job performance that includes effort and ability, as Walker, Churchill, and Ford (1977) advocate and as

---

[11] We estimated a model that combined the three role perception variables into a single "job tension" factor. We found a marginally significant effect on job satisfaction ($\delta = -.100, p < .1$) and an insignificant effect on effort ($\delta = -.096, p = .28$). In addition, we allowed for direct effects of job factors and role perception variables on store performance and found no significant effect and no change in the parameter estimates of interest. However, these factors could potentially moderate the effect of effort on job performance. Such a test is left to further research.

Lusch and Serpkenci (1990) use. We then estimate this same model (without effort) with the "narrow" measure of job performance we used in our original model. This replicates Bagozzi's (1978) approach. We expect the first alternate model to replicate the insignificant effect of job performance on job satisfaction that Lusch and Serpkenci (1990) report: The negative and the positive elements in the job performance measure should cancel each other out. By completely eliminating effort, the second version should yield a somewhat larger effect of job performance on job satisfaction, but it should still be substantially smaller than the one we obtained from the fully specified four-equation model.

The estimation results shown in the second and third columns of Table 4 support these conjectures. (To facilitate the comparison, we repeated the relevant results from Table 3 in the first column of Table 4.) Using an aggregate measure of job performance leads to an insignificant effect of job performance on job satisfaction ($\beta = .023, p = .76$). This result is equivalent to the insignificant effect that Lusch and Serpkenci (1990) report. The use of the narrow measure of job performance leads to a positive but much smaller effect ($\beta = .136, p < .05$).[12] These results strongly support H_2 and indicate that failing to control properly for the effect of effort on job satisfaction leads to results that are biased toward insignificance.[13]

Note that the effects of different job factors on job performance and job satisfaction are different in the alternate models. In particular, when we exclude effort from the job performance measure, the results show a negative and significant effect of job attractiveness and a negative and marginally significant effect of job autonomy on job perfor-

---

[12] The second alternate model is a nested version of the full model. As a result, we can statistically test the implied restrictions using a chi-square test. The alternate model can be rejected with a high degree of significance ($\chi^2 = 21.4$, degree of freedom = 1).

[13] We also estimated a model similar to that of Brown and Peterson (1994), which excluded job characteristics, especially the compensation variables. In this case, we found a positive effect of effort.

TABLE 4
Estimation Results for Alternate Models

		Job Satisfaction Models[a]	
Parameter	**Relevant Results from Table 3**	**Aggregate Job Performance Measure**	**Narrow Job Performance Measure**
Job Satisfaction			
Constant	–.054 (.051)	–.044 (.066)	–.072 (.077)
Effort	–.269*** (.061)		
Job performance	.286*** (.054)	.023 (.076)	.136** (.062)
Compensation	.168*** (.057)	.299*** (.085)	.239*** (.086)
Profit sharing	.246*** (.080)	.163 (.121)	.166 (.119)
Job autonomy	.096* (.061)	.285*** (.084)	.380*** (.079)
Job attractiveness	.476*** (053)	.304*** (.092)	.357*** (.092)
Supervisory feedback	–.128 (.060)	–.081 (.091)	–.099 (.089)
Social climate	.176*** (.056)	.168** (.083)	.190** (.083)
Role ambiguity	–.026 (.060)	.019 (.097)	–.026 (.090)
Role conflict	–.047 (.031)	.026 (.073)	–.009 (.054)
Role overload	–.083 (.045)	.057 (.081)	–.040 (.080)
Effort/Job Performance[b]			
Constant	.078 (.213)	–.190 (.331)	–.630* (.376)
Compensation	–.001 (.087)	.090 (.075)	.186* (.086)
Profit sharing	.183* (.109)	.289*** (.111)	.134 (.126)
Job autonomy	.249*** (.079)	.166** (.074)	–.101 (.081)
Job attractiveness	.020 (.092)	–.088 (.081)	–.195** (.081)
Supervisory feedback	140* (.080)	.129* (.078)	.052 (.088)
Social climate	–.204** (.083)	–.213*** (.072)	–.107 (.083)
Role ambiguity	.093 (.091)	.0711 (.063)	–.106 (.090)
Role conflict	–.024 (.054)	–.028 (.072)	.045 (.055)
Role overload	–.119* (.070)	–.096 (.071)	–.156* (.080)
χ^2 (d.f.)		12.9*** (1)	21.4*** (1)

*$p < .10$ (two-tailed t-test).
**$p < .05$ (two-tailed t-test).
***$p < .01$ (two-tailed t-test).
[a] We left the control variables as in the original model.
[b] Because the two job satisfaction models did not include effort, the various job characteristic and role perception variables were directly linked to job performance.
Notes: Standard errors are indicated in parentheses.

mance. Again, these results suggest that omitting effort leads to biased estimates.

Discussion

This study was motivated by inconsistencies across different studies in the job satisfaction research, including limited and contradictory evidence that high levels of effort are costly for employees. An examination of the literature suggests that the inconsistencies most likely arise from differences in construct definitions and a failure to account for effort properly. Effort and compensation are the most important factors in a vertical work relationship, yet few job satisfaction studies account for both as separate constructs. Furthermore, in marketing, the majority of studies examine the job satisfaction of salespeople (i.e., a context in which compensation is frequently related to job performance and, thus, effort). The failure to account for effort here (and, thus, the effort–compensation relationship) also biases estimates of the relationship between compensation and job satisfaction. In addition, we were interested in understanding the effects of corporatewide profit-sharing plans, which is a relatively unresearched element of the compensation package in many organizations.

The Relationship Between Job Performance and Job Satisfaction

The empirical analysis in this article uses data collected for a typical job satisfaction study, but the model we propose incorporates the basic constructs from an agency relationship and the key antecedents of job satisfaction. In contrast to previous models, this framework accounts for both the direct and the indirect effects of effort (through job performance) on job satisfaction and leads to conclusions that are significantly different.

First, we find a significant, positive effect of job performance on job satisfaction. This contrasts with Lusch and Serpkenci's (1990) reported negative (but insignificant) effect of job performance, using the same data set. We show that not accounting for effort biases the estimated effect of job performance on job satisfaction, and this can lead to incorrect conclusions. Second, by changing the model and construct measures to be consistent with prior studies, we replicate small or insignificant effects of job performance on job satisfaction and a positive effect of effort on job satisfaction, as reported in the existing literature. This supports the argument that problems with construct definition can explain inconsistent findings in the job satisfaction literature. This finding also challenges the now widely accepted view about the relationship between job performance and job satisfaction in organizational psychology and the sales force literature that the two constructs do not affect each other. In particular, decades of research have failed to find a significant or consistent link between job performance and job satisfaction (Brown and Peterson 1993; Iaffaldano and Muchinsky 1985).

A positive effect of job performance on job satisfaction also has important implications for a firm that wants to motivate and retain talented employees. It implies that actions to increase job performance can also increase employees' job satisfaction. As a result, benefits such as reduced turnover and less absenteeism (a result of higher job satisfaction) may be useful for justifying the cost of a policy, even when that policy is directed primarily toward improving job performance.

Effort, Job Performance, and Job Satisfaction

Because job satisfaction is a proxy for utility (Clark and Oswald 1996; Friedman 1978), a negative relationship between manager effort and job satisfaction follows from the principal–agent framework. Nevertheless, the relationship has proved elusive to empirical validation. Recognizing the different possible paths through which an employee's effort may affect his or her satisfaction enables us to confirm a negative relationship empirically. This finding, combined with the positive effect of job performance on store performance, is consistent with the economic view that there is an inherent conflict of interest in vertical relationships. When accounting for effort, we also find relatively weak effects of the role perception variables. Factors such as role ambiguity and role conflict are considered key determinants of job satisfaction in the sales force literature (Brown and Peterson 1993). Because these factors capture work-related stress, it is possible that they captured the omitted effect of effort.

The positive effect of job performance on job satisfaction also has important implications for research based on agency theory. In a standard agency model (e.g., Lal, Outland, and Staelin 1994), it is assumed that job performance can increase job satisfaction only indirectly (through factors such as compensation or promotion). Our research demonstrates that this standard assumption of agency theory may not be justified. The omission of job performance's effect on job satisfaction has the potential to bias the effect of effort on job satisfaction toward insignificance.[14] For example, when the effect of job performance is eliminated from the job satisfaction equation (Equation 4), we obtain a negative but insignificant effect of effort on job satisfaction ($\hat{\beta} = -.031$, $p = .24$).[15]

Corporate Profit-Sharing Plan

In general, the literature distinguishes between two types of income that employees can earn: fixed compensation, which is paid to the employee independent of output, and variable pay, which is paid to the employee contingent on output. In our data, there is a third type of compensation: payments that are based on overall corporate performance. Because corporate performance is not known in advance, the expected payment to the employee under this plan is uncertain. In contrast to variable pay, which is directly contingent on an employee's output, the ability of an employee in a large organization, such as a retail chain, to affect the level of the expected payoff from a corporate-level plan is also low. The findings show that this type of compensation has some incentive pay–type effects: It has positive effects on both effort and job satisfaction. This holds even though increased effort by a store manager is unlikely to affect the ultimate payout from the plan significantly. This is useful because many firms would like to provide incentives to their employees (with performance-based pay) but cannot because the impact of an individual employee on output cannot be isolated. Most employees in marketing positions have a relatively small direct effect on the outcome of their brand, business unit, or customer service department. In these conditions, profit-sharing plans based on the performance of the overall organization provide an opportunity to obtain an effect similar to that of performance-based pay for the sales force. Thus, although data limitations prevent us from conducting a detailed analysis of why corporate profit-sharing plans act the way they do, our results indicate that these plans can be of significant value in organizations.

---

[14] This is analogous to the way that the omission of effort's effect on job satisfaction biases the effect of job performance toward insignificance.

[15] Clark (1997) does not include job performance and also finds a relatively weak negative effect of effort.

Summary and Further Research

This article attempts to clarify ambiguities in the literature about the relationships among effort, job performance, and job satisfaction. Using a model with variables that are important for both agency theory and organizational psychology, we find a negative, direct effect of effort and a positive, direct effect of job performance on job satisfaction. We show that conflicting findings in the literature are a result of inconsistency in both the measurement and the definition of constructs across studies that do not fully account for all the relationships among constructs. This suggests that some results in the sales force literature regarding job satisfaction, such as the strong effect of role perception variables, should be reexamined. This article also demonstrates important omitted variables biases that can arise in the empirical analysis of work relationships.

The most obvious extension is to generalize the findings to agency relationships in which the incentives in the compensation packages are stronger (i.e., there are variable components in the compensation that are directly related to the agent's output). Restricting the analysis to a single retailer provides a benefit by reducing the potential number of confounding factors. We also identified an interesting finding about corporatewide profit-sharing plans. However, we cannot determine its relative effectiveness with the available data. An interesting issue for further investigation is to determine whether and how different types of employees respond to pay that is based on overall corporate performance.

Similar to any empirical study of complex organizational issues, this study is subject to several cautions. For some of our measures, we use cross-sectional data, and this limits our ability to control fully for certain unobserved effects. In addition, it limits our ability to test the direction of causality. To the extent possible, we make use of covariates and test for reverse relationships. We have also exercised care to reduce potential problems pertaining to measurement errors, sampling bias, and incomplete model specification. That being said, the study has two methodological aspects that differentiate it from other studies. First, we use individual-level data to test an individual-level model, as Lal, Outland, and Staelin (1994) suggest. Second, we use multiple sources of information to measure the various constructs, thus reducing the problem of common method bias and common source bias. This enables us to shed new light on the determinants of job satisfaction and provide insights that are important both for researchers in this domain and for organizations.

APPENDIX
Scale Items to Measure Model Constructs

Compensation

I am paid fairly for the work I do.[a]

My pay is better than that for similar jobs in other firms.[a]

Salary and wage increases are given to those who do a good job.[a]

Do you feel your pay is as high in comparison with what others get for similar work in other companies?[b]

Job Attractiveness

After a day's work, I really feel like I have accomplished something.[a]

I am in a "dead-end" job. (R)[a]

I am often bored with my job. (R)[a]

Do you find your work challenging, exciting, and giving you a sense of accomplishment?[b]

Role Conflict

Having to decide things that affect the lives of individuals that you know.[c]

Feeling that you may not be liked and accepted by the people you work with.[c]

Feeling that you do things on the job that are against your better judgment.[c]

Role Overload

Feeling that you have heavy a workload, one that you can't possibly finish during an ordinary workday.[c]

Thinking that the amount of work you have to do interferes with how well it gets done.[c]

Feeling that your job tends to interfere with your family life.[c]

APPENDIX
Continued

Job Autonomy

I have a lot to say about how to do my job.[a]

How satisfied are you with the amount of control you have in your work?[f]

Supervisory Feedback

Management is quick to criticize poor performance. (R)[a]

No one ever says "you've done a good job." (R)[a]

Do you feel you do not know what your supervisor thinks of you, how he evaluates your performance? (R)[b]

Do you feel free to offer suggestions concerning policies and procedures affecting your operation?[b]

Social Climate

Working for Winn's is like being a part of a family.[a]

The people here are proud to work for Winn's.[a]

Role Ambiguity

Being unclear on just what the scope and responsibilities of your job are.[c]

Feeling that you are not fully qualified to handle your job.[c]

The fact that you can't get enough information to carry out your job.[c]

Not knowing just what the people you work with expect of you.[c]

Effort

Takes responsibility in his work.[d]

Readily assumes responsibility.[d]

Makes an effort to improve his managerial skills.[d]

Works long hours when necessary.[d]

Level of motivation.[e]

Ability

Can cope with pressure or strain on the job.[d]

Decision-making ability.[e]

Knowledge of trade area.[e]

Tolerance for pressure.[e]

Job Performance

Fails to meet to target goals set for the store. (R)[d]

Achieving the target sales volume.[e]

Achieving the target gross profit.[e]

Job Satisfaction

How satisfied are you with your general work situation?[f]

Would you advise a friend looking for a new job to take one similar to yours?[b]

I just hate to get up in the morning to go to work.[a]

Source: Based on instruments that Lusch and Serpkenci (1990) use.
[a]Measured on a five-point Likert scale, ranging from "strongly agree" to "strongly disagree"
[b]Measured on a six-point Likert scale, ranging from "definitely yes" to "definitely no."
[c]Measured on a five-point Likert scale, ranging from "never bothered" to "bothered nearly all the time."
[d]Measured on a five-point Likert scale, ranging from "strongly disagree" to "strongly agree."
[e]Measured on a seven-point Likert scale, ranging from "very satisfied" to "very dissatisfied."
[f]Measured on a six-point Likert scale, ranging from "extremely satisfied" to "extremely dissatisfied."
Notes: R = reverse coded.

REFERENCES

Anderson, Erin and Richard L. Oliver (1987), "Perspective on Behavior-Based Versus Outcome-Based Sales Force Control Systems," *Journal of Marketing*, 51 (October), 76–88.

Bagozzi, Richard P. (1978), "Sales Force Performance and Satisfaction as a Function of Individual Difference, Interpersonal, and Situational Factors," *Journal of Marketing Research*, 15 (November), 517–31.

——— (1980), "Performance and Satisfaction in an Industrial Sales Force: An Examination of their Antecedents and Simultaneity," *Journal of Marketing*, 44 (Spring), 65–77.

Behrman, Douglas N. and William D. Perreault (1984), "A Role Stress Model of the Performance and Satisfaction of Industrial Salespersons," *Journal of Marketing*, 48 (Fall), 9–21.

Boyt, Thomas E., Robert F. Lusch, and Gillian Naylor (2001), "The Role of Professionalism in Determining Job Satisfaction in Professional Services," *Journal of Service Research*, 3 (May), 321–30.

Brown, Steven P. and Robert A. Peterson (1993), "Antecedents and Consequences of Salesperson Job Satisfaction: Meta-Analysis and Assessment of Causal Effects," *Journal of Marketing Research*, 30 (February), 63–77

——— and ——— (1994), "The Effect of Effort on Sales Performance and Job Satisfaction," *Journal of Marketing*, 58 (April), 70–80.

Campbell, John P. and Robert D. Pritchard (1976), "Motivation Theory in Industrial and Organizational Psychology," in *Handbook of Industrial and Organizational Psychology*, M. Dunnette, ed. Chicago: Rand McNally, 63–130.

Clark, Andrew E. (1997), "Job Satisfaction and Gender: Why Are Women So Happy at Work?" *Labour Economics*, 4 (4), 341–72

——— and Andrew J. Oswald (1996), "Satisfaction and Comparison Income," *Journal of Public Economics*, 61 (3), 359–81.

Deci, Edward L. (1971), "Effects of Externally Mediated Rewards on Intrinsic Motivation," *Journal of Personality and Social Psychology*, 18 (1), 105–115.

——— and Richard M. Ryan (1985), *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum.

Dunne, Patrick and Robert F. Lusch (1999), *Retailing*. Fort Worth, TX: Dryden Press.

Fried, Yitzhak and Gerald R. Ferris (1986), "The Dimensionality of Job Characteristics: Some Neglected Issues," *Journal of Applied Psychology*, 71 (3), 419–26.

Friedman, Richard B. (1978), "Job Satisfaction as an Economic Variable," *American Economic Review*, 68 (May), 135–41.

Greene, William H. (1991), *Econometric Analysis*. New York: Macmillan.

Hackman, J. Richard and Greg R. Oldham (1975), "Development of the Job Diagnostic Survey," *Journal of Applied Psychology*, 60 (2), 159–70.

Hart, Oliver D. and Bengt Holmstrom (1987), "The Theory of Contracts," in *Advances in Economic Theory, Fifth World Congress*, T. Bewley, ed. New York: Cambridge University Press.

Holmstrom, Bengt (1979), "Moral Hazard and Observability," *Bell Journal of Economics*, 10 (1), 74–91.

Iaffaldano, Michelle T. and Paul M Muchinsky (1985), "Job Satisfaction and Job Performance: A Meta-Analysis," *Psychological Bulletin*, 97 (2), 251–73.

Judge, Timothy A., Joyce E. Bono, and Edwin A. Locke (2000), "Personality and Job Satisfaction: The Mediating Role of Job Characteristics," *Journal of Applied Psychology*, 85 (2), 237–49.

Lahiri, Kajal and Peter Schmidt (1978), "On the Estimation of Triangular Structural Systems," *Econometrica*, 46 (5), 1217–21.

Lal, Rajiv, Don Outland, and Richard Staelin (1994), "Salesforce Compensation Plans: An Empirical Test of the Agency Theory Framework," *Marketing Letters*, 5 (2), 117–30.

Lawler, Edward E. and Lyman W. Porter (1967), "The Effect of Performance on Job Satisfaction," *Industrial Relations*, 7, 20–28.

Locke, Edwin A. (1969), "The Nature and Causes of Job Satisfaction," in *Handbook of Industrial and Organizational Psychology*, M.D. Dunnette, ed. Chicago: Rand McNally.

Lusch, Robert F. and Ray R. Serpkenci (1990), "Personal Differences, Job Tension, Job Outcomes, and Store Performance: A Study of Retail Store Managers," *Journal of Marketing*, 54 (January), 85–101.

Ostroff, Cherri (1992), "The Relationship Between Satisfaction, Attitudes and Performance: An Organizational Level Analysis," *Journal of Applied Psychology*, 77 (6), 963–74.

Porter, Lyman W. and Edward E. Lawler (1968), *Managerial Attitudes and Performance*. Homewood, IL: Richard D. Irwin.

Rizzo, John R., Robert J. House, and Sidney I. Lirtzman (1970), "Role Conflict and Ambiguity in Complex Organizations," *Administrative Science Quarterly*, 15 (June) 150–63.

Ryan, Richard M., Valerie Mims, and Richard Koestner (1983), "Relation of Reward Contingency and Interpersonal Context to Intrinsic Motivation: A Review and Test Using Cognitive Evaluation Theory," *Journal of Personality and Social Psychology*, 45 (4), 736–50.

Scarpello, V. and John P. Campbell (1983), "Job Satisfaction: Are All the Parts There?" *Personnel Psychology*, 36 (3), 577–600.

Smith, P.C., L.M. Kendall, and C.L. Hulin (1969), *The Measurement of Satisfaction in Work and Retirement*. Chicago: Rand McNally.

Walker, Orville C., Gilbert A. Churchill Jr., and Neil M. Ford (1977), "Motivation and Performance in Industrial Selling: Present Knowledge and Needed Research," *Journal of Marketing Research*, 14 (May), 156–68.

TEXERE – The Intersection of Insight & Innovation **THOMSON**

The **American Marketing Association** is proud to have a strategic alliance with Texere, a product line of Thomson Professional Business, to deliver a new generation of professional marketing resources to help you better meet the demands of today's challenging marketing arena. Our vision for the new AMA Bookstore is to bring you a full library of the most cutting-edge resources for marketing professionals filled with practical marketing solutions and insight into innovative marketing strategies.

NEW Books from the AMA Bookstore

Creating Do-It-Yourself Customers: How Great Customer Experiences Build Great Companies
Self-checkout, self-customization, and other co-production efforts are on the rise. But it's a boon to business only if customers do their share. This book helps readers recognize situations where high-tech or human innovation can bolster customer involvement.
Peter Honebein/Roy Cammarano
0-324-31154-0 — $29.95 / AMA Members $23.95

Strategic Sales Leadership: BREAKthrough Thinking for BREAKthrough Results
While there are numerous cookbook approaches to sales management, very little is available on training executive sales managers how to think and manage their sales teams strategically. Until now. This book equips sales executives with the tools to develop an effective action plan to improve sales organization results
The Sales Educators
1-58799-203-5 — $39.95 / AMA Members $31.95

Ten Key Customer Insights: Unlocking the Mind of the Market
While companies tout "Customer is King," the reality is declining market share, new product failure, and financial woes. This book helps readers develop marketing strategies and tactics to attract and retain more customers, grow market share, increase the productivity of their marketing efforts, and boost profitability
Robert Schieffer
1-58799-206-X — $29.95 / AMA Members $23.95

Part I: Marketing Scales Handbook Volume IV: A Compilation of Multi-Item Measures for Consumer Behavior & Advertising
This updated version provides all new or new uses of previously developed scales in consumer behavior and advertising that appeared between 1998 and 2001 The authors present 654 scales with a description, the origin, reliability, validity, and more. This is Part I of a two part set for Volume IV.
Bruner/Hensel/James
1-58799-205-1 — $98.95 / AMA Members $80.00

The Case for B2B Branding: Pulling Away from the Business-to-Business Pack
Intel and Performance. 3M and Innovation. Companies with the most highly developed brand images can be associated with a single attribute. Though much has been written about branding at the consumer level, B2B branding is a completely different game. This enlightening new book delivers the secrets to success
Bob Lamons
0-324-39865-4 — $29.95 / AMA Members $23.95

One Customer, Divisible: Linking Customer Insight to Loyalty and Advocacy Behavior
Discover how one book could teach you how to optimize your customer relationships. *One Customer* goes beyond traditional thinking approaches to CRM to emphasize a key leveraging factor largely missed by most marketers thus far: using individual customer data to provide the highest level of perceived value
Michael Lowenstein
0-324-30129-4 — $29.95 / AMA Members $23.95

Data-Driven Business Models
Develop and run your business more effectively using metrics to measure performance Weber shows how to model and continuously reinvent your business to meet the needs of current customers, and identify and capitalize on new profitable opportunities.
Alan Weber
0-324-22233-5 — $49.95 / AMA Members $39 95

Creating the Marketing Experience: New Strategies for Building Relationships with Your Target Market
Companies are concerned about the increased difficulty in developing marketing vehicles that cut through the clutter and make a positive impact. This book helps marketers create the contexts in which prospects can experience the brand on their own terms.
Joe Marconi
0-324-20544-9 — $54.95 / AMA Members $43.95

Escaping the Black Hole: Minimizing the Damage from the Marketing-Sales Disconnect
A fourth of all B2B marketing and sales resources are wasted on misguided initiatives and activities This enlightening book aligns marketing and sales organizations and reveals the secrets to bridging the gap, cutting the clutter, and eliminating waste.
Robert Schmonsees
0-324-30125-1 — $29.95 / AMA Members $23 95

Business to Business Marketing Research
Spotlighting new techniques related to focus groups and in-depth interviews as well as the impact of data mining and other technology-driven types of research—this new AMA book covers topics that aren't covered anywhere else!
Martin P. Block / Tamara S. Block
0-324-22230-0 — $69.95 / AMA Members $55.95

To Order: **1-800-842-3636** http://www.marketingpower.com/amabooks

On the global**EDGE**?

Is your business really prepared for international expansion? Do you want to know how you can go global? Well....find out from your personal online consultant!

CORE™ is a self-assessment tool that will allow you to determine your companyís readiness to expand its operations internationally and ascertain its ability to export a particular product.

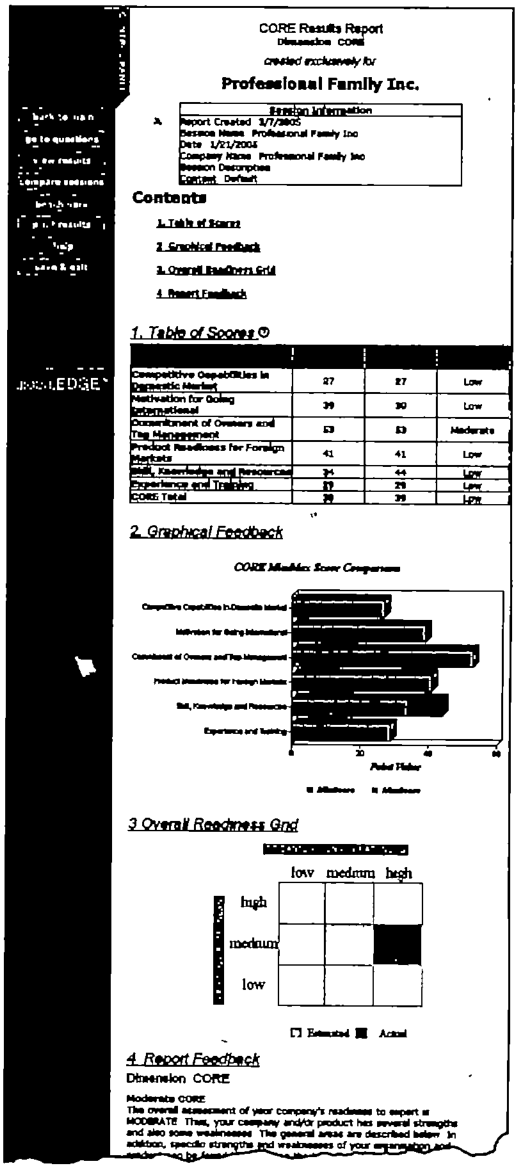
Partner™ is an interactive tool that will assist you in evaluating and comparing a variety of potential international venture partners.

Distributor™ is designed to aid your company with identifying the best distributor or agent to use (specific to your product and market characteristics) when moving into new markets.

Freight™ evaluates the compatibility of international freight forwarders in respect to your business' type and volume of freighting needs.

GLOBE™ (under development) will help you assess the degree of globalization potential of your organization based on your business' structure, culture, management processes, and strategy.

DEVELOPED BY THE CENTER FOR INTERNATIONAL BUSINESS EDUCATION AND RESEARCH AT MICHIGAN STATE UNIVERSITY



CORE Results Report
Dimension CORE
created exclusively for
Professional Family Inc.

Session Information
Report Created 1/7/2005
Session Name Professional Family Inc
Date 1/21/2005
Company Name Professional Family Inc
Session Description
Content Default

Contents

1. Table of Scores
2. Graphical Feedback
3. Overall Readiness Grid
4. Report Feedback

1. Table of Scores

Competitive Capabilities in Domestic Market	27	27	Low
Motivation for Going International	39	39	Low
Commitment of Owners and Top Management	53	53	Moderate
Product Readiness for Foreign Markets	41	41	Low
Skill, Knowledge and Resources	34	44	Low
Experience and Training	29	29	Low
CORE Total	39	39	Low

2. Graphical Feedback

CORE Min/Max Score Comparison
Competitive Capabilities in Domestic Market
Motivation for Going International
Commitment of Owners and Top Management
Product Readiness for Foreign Markets
Skill, Knowledge and Resources
Experience and Training
0
20
40
60
Point Value
Minscore
Maxscore

3 Overall Readiness Grid

4 Report Feedback

Dimension CORE

Moderate CORE

The overall assessment of your company's readiness to export is MODERATE. Thus, your company and/or product has several strengths and also some weaknesses. The general areas are described below. In addition, specific strengths and weaknesses of your organization and

The Diagnostic Tools are web-based software programs that incorporate the collective knowledge and experiences of seasoned international business executives and empirical research findings from international business operations of the past 20 years.

Diagnostic Tools globalEDGE.msu.edu

http://globalEDGE.msu.edu/DiagTools

global**EDGE**™

YOUR SOURCE FOR global business knowledge

globalEDGE is the #1 web resource for 'international business' information in the world, based on rankings from Google as of May 1, 2005

Journal of Marketing

pril 2006
olume 70, Number 2

A Quarterly Publication of the
American Marketing Association

The Emergence of Dominant Designs 1 Raji Srinivasan, Gary L. Lilien, & Arvind Rangaswamy
Drivers of Brand Extension Success 18 Franziska Völckner & Henrik Sattler
Crossover Dreams: Consumer Responses to Ethnic-Oriented Products 35 Sonya A. Grier, Anne M. Brumbaugh, & Corliss G. Thornton
Brand Dilution: When Do New Brands Hurt Existing Brands? 52 Chris Pullig, Carolyn J. Simmons, & Richard G. Netemeyer
Cross-Functional "Coopetition": The Simultaneous Role of Cooperation and Competition Within Firms 67 Xueming Luo, Rebecca J. Slotegraaf, & Xing Pan
Consumer Contamination: How Consumers React to Products Touched by Others 81 Jennifer J. Argo, Darren W. Dahl, & Andrea C. Morales
Ratios in Proportion: What Should the Shape of the Package Be? 95 Priya Raghubir & Eric A. Greenleaf
Stuck in the Past: Why Managers Persist with New Product Failures 108 Eyal Biyalogorsky, William Boulding, & Richard Staelin
The Evolution of Loyalty Intentions 122 Michael D. Johnson, Andreas Herrmann, & Frank Huber
Converting Web Site Visitors into Buyers: How Web Site Investment Increases Consumer Trusting Beliefs and Online Purchase Intentions 133 Ann E. Schlosser, Tiffany Barnett White, & Susan M. Lloyd

AMERICAN
MARKETING
ASSOCIATION

"Covers brand management, brand extension, name changes, brand revitalization, and global brand management. Recommended." —CHOICE magazine

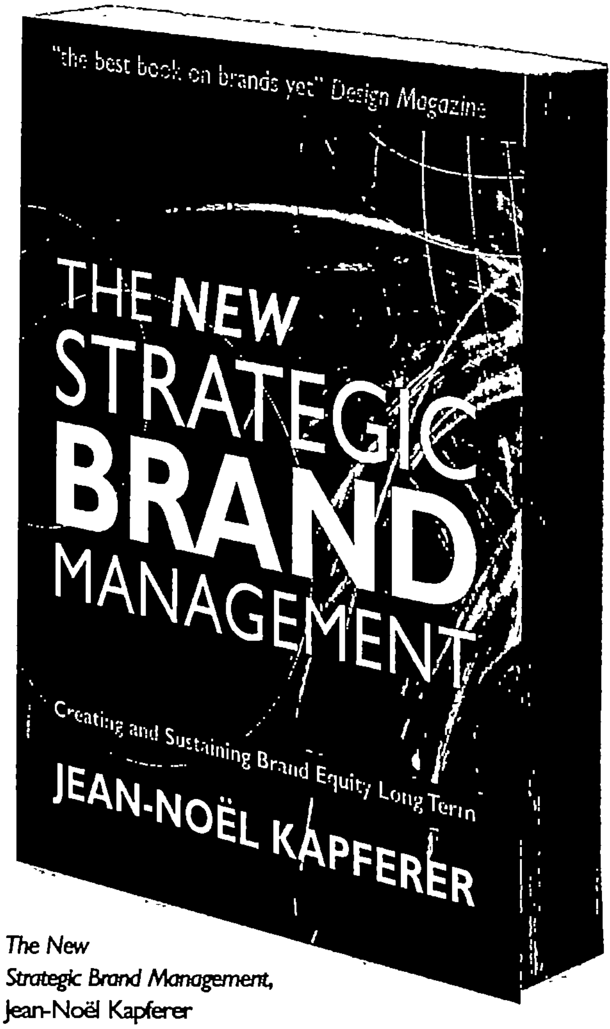
"Creatively written…challenging in its thinking yet interesting and accessible." —Media Week

THE MOST INNOVATIVE, THOROUGH AND RELEVANT BOOK ON BRAND MANAGEMENT

Innovative: proposes new models, new concepts, new benchmarks for brand and business building, all absent from other books on brands. New approaches to brand identity and positioning, brand architectures, brand portfolios and brand globalization.

Thorough: analyzes in depth all facets of brand management, with practical models and advice.

Relevant: no other book addresses the key issues of brand management today including: **Building brands in saturated markets • What innovation for what brands? • Managing retail brands • Activating the brand • Energizing the brand value chain • Globalizing existing brands • Brand extensions.**


"the best book on brands yet" Design Magazine
THE NEW
STRATEGIC
BRAND
MANAGEMENT
Creating and Sustaining Brand Equity Long Term
JEAN-NOËL KAPFERER

The New Strategic Brand Management,
Jean-Noël Kapferer
Paper, 0 7494 4283 2, $55.00
Published by Kogan Page

About the author: **JN Kapferer** (HEC, Ph.D Northwestern University) is considered one of the top three branding experts in the world. His style of analysis uniquely blends key research developments with very practical concerns. He has experience as a consultant for European, Asian and American companies and teaches at advanced executive development seminars worldwide.

KOGAN PAGE US

For more details, full table of contents, to download a sample chapter and order, visit
www.kogan-page.co.uk/kapferer or
www.styluspub.com
Inspection copies available

Journal of Marketing

Vol. 70, No. 2

AMERICAN MARKETING ASSOCIATION

EDITOR
Roland T. Rust
Robert H. Smith School of Business
3451 Van Munching Hall
University of Maryland
College Park, MD 20742-1815

ASSISTANT TO THE EDITOR
Marti Tiedeman
E-mail: jom@ama.org
Tel: (312) 542-9024

Millicent Locke
University of Maryland

WEB EDITOR
Siva K. Balasubramanian
Southern Illinois University

MARKETING LITERATURE REVIEW
Myron Leonard
Western Carolina University

BOOK REVIEWS
Ramarao Desiraju
University of Central Florida

VICE PRESIDENT—PUBLICATIONS
Russell S. Winer
New York University

CHIEF EXECUTIVE OFFICER
Dennis Dunlap

GROUP PUBLISHER
Jack Hollfelder

EDITORIAL STAFF
Phone: (312) 542-9000
Fax: (312) 542-9001

MANAGING EDITOR
Francesca Van Gorp Cooley

TECHNICAL EDITOR
Andy Seagram

GRAPHICS EDITOR
Mary Loye

Advertising Staff
Fax: (312) 922-3763

Richard Ballschmiede, Advertising
(312) 542-9076

Sally Schmitz, Production
(312) 542-9038

JM Web site:
Submissions: www.marketingjournals.org/JM
Main page: www.marketingpower.com/JM

The *Journal of Marketing* (ISSN 0022-2429) is published quarterly (January, April, July, and October). *Annual Subscription Rates:* Subscriptions are $90 for individuals in the United States and Canada and $115 in other countries. Library, school, and business subscriptions are $225 in the United States and Canada and $250 in other countries Subscriptions for AMA members are $53. Prices are for one year. Single copy rate: $50 individual, $50 corporate in the United States and Canada; add $5 in other countries. Prepayment is required in U.S. dollars or the equivalent.

Published by the Publications Group of the American Marketing Association, 311 S Wacker Dr, Suite 5800, Chicago, IL 60606-2266. Periodicals postage paid at Chicago, IL and additional mailing offices. POSTMASTER· Send address changes to the *Journal of Marketing*, 311 S Wacker Dr., Suite 5800, Chicago, IL 60606-2266, USA.

•**MANUSCRIPTS**: Electronic submissions must be made via the *JM* Web site at http://www marketingjournals.org. Paper submissions are not accepted. Editorial policies, review board members, Tables of Contents for *JM*, and so on, are available on the *JM* Web site at http://www.marketingpower.com/jm.

•**ADVERTISERS AND ADVERTISING AGENCIES** assume liability for all content (including text, representations, and illustrations) of advertisements published, and also assume responsibility for any claims arising therefrom made against the publisher. The right is reserved to reject any advertisement.

•**SUBSCRIPTIONS***:* Members Send to American Marketing Association, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266 (1-800-AMA-1150 or 312/542-9000), or e-mail to info@ama.org. Nonmembers: Send to American Marketing Association, P.O. Box 11806, Birmingham, AL 35202, 1-800-633-4931, 205-995-1567, Fax. 205-995-1588, e-mail: amasubs@ebsco.com. Canadian residents add 7% GST (GST registration # 12747852).

•**CHANGE OF ADDRESS:** Send new address, address label from *JM*, and zip code to American Marketing Association Allow 6 to 8 weeks.

•**PERMISSIONS AND REPRINT INFORMATION***:* Where necessary, permission is granted by the copyright owner for those registered with the Copyright Clearance Center (CCC), 222 Rosewood Drive, Danvers, MA 01923, to photocopy any article herein for a flat fee Contact the CCC with the ISSN (0022-2429), volume, and first and last page numbers of each article copied Copying for other than personal or internal reference without express permission of the American Marketing Association Publications Group is prohibited. Address requests for customized bulk reprints to Kristen Ramsey, Valeo IP, 651-582-3851 or 1-800-217-7874.

The views expressed in *JM* are not necessarily those of the editorial staff or the publisher. **All articles in this journal, unless otherwise noted, have undergone a blind review process.**

Copyright © 2006, American Marketing Association All rights reserved.

Editorial Review Board

ROHINI AHLUWALIA
University of Minnesota

KUSUM L AILAWADI
Dartmouth College

TIM AMBLER
London Business School

EUGENE W ANDERSON
University of Michigan

TOR WALLIN ANDREASSEN
Norwegian School of Management

ASIM ANSARI
Columbia University

ERIC J. ARNOULD
University of Arizona

BARRY L BAYUS
University of North Carolina

PAUL D. BERGER
Boston University

PAUL N BLOOM
University of North Carolina

LISA E BOLTON
University of Pennsylvania

RUTH N. BOLTON
Arizona State University

ROBERT BORDLEY
General Motors Corporation

DOUGLAS BOWMAN
Emory University

RODERICK J BRODIE
University of Auckland

DIPANKAR CHAKRAVARTI
University of Colorado

PIERRE CHANDON
INSEAD

PRADEEP K. CHINTAGUNTA
University of Chicago

KEVIN CLANCY
Copernicus Marketing Consulting

PETER J DANAHER
University of Auckland

JOHN A. DEIGHTON
Harvard University

MARNIK G. DEKIMPE
Catholic University Leuven

WAYNE S. DESARBO
Pennsylvania State University

RAVI DHAR
Yale University

NAVEEN DONTHU
Georgia State University

TULIN ERDEM
University of California, Berkeley

PETER S. FADER
University of Pennsylvania

GARY L FRAZIER
University of Southern California

PETER N. GOLDER
New York University

RAJDEEP GREWAL
Pennsylvania State University

ABBIE GRIFFIN
University of Illinois, Urbana-Champaign

SUNIL GUPTA
Columbia University

JOHN R HAUSER
Massachusetts Institute of Technology

JAN B. HEIDE
University of Wisconsin, Madison

CHRISTIAN HOMBURG
University of Mannheim

WAYNE D. HOYER
University of Texas at Austin

G TOMAS M. HULT
Michigan State University

J JEFFREY INMAN
University of Pittsburgh

GANESH IYER
University of California, Berkeley

CHRIS JANISZEWSKI
University of Florida

SANDY D JAP
Emory University

BERNARD J JAWORSKI
Monitor Group

MICHAEL D JOHNSON
University of Michigan

P.K. KANNAN
University of Maryland

TIMOTHY L KEININGHAM
Ipsos Loyalty

AMNA KIRMANI
University of Maryland

AJAY K KOHLI
Emory University

ROBERT V KOZINETS
York University

V KUMAR
University of Connecticut

DONALD R. LEHMANN
Columbia University

KATHERINE N. LEMON
Boston College

MARY FRANCES LUCE
Duke University

ROBERT F. LUSCH
University of Arizona

JOHN G. LYNCH JR.
Duke University

DEBORAH J. MACINNIS
University of Southern California

VIJAY MAHAJAN
University of Texas at Austin

CARL F. MELA
Duke University

VIKAS MITTAL
University of Pittsburgh

CHRISTINE MOORMAN
Duke University

PAGE MOREAU
University of Colorado

SCOTT A NESLIN
Dartmouth College

STEPHEN M. NOWLIS
Arizona State University

JOSEPH C NUNES
University of Southern California

A PARASURAMAN
University of Miami

WILLIAM D PERREAULT
University of North Carolina

F G M. PIETERS
Tilburg University

JAIDEEP PRABHU
Imperial College

ARVIND RANGASWAMY
Pennsylvania State University

BRIAN T. RATCHFORD
University of Texas at Dallas

AMERICUS REED
University of Pennsylvania

WERNER REINARTZ
INSEAD

STEVEN M. SHUGAN
University of Florida

ITAMAR SIMONSON
Stanford University

DILIP SOMAN
University of Toronto

JOYDEEP SRIVASTAVA
University of Maryland

JAN-BENEDICT E.M. STEENKAMP
Tilburg University

DAVID W. STEWART
University of Southern California

STEFAN STREMERSCH
Erasmus University Rotterdam

BAOHUNG SUN
Carnegie Mellon University

GERARD J. TELLIS
University of Southern California

STIJN VAN OSSELAER
Erasmus University Rotterdam

P. RAJAN VARADARAJAN
Texas A&M University

STEPHEN L VARGO
University of Hawaii, Manoa

PETER VERHOEF
University of Groningen

MICHEL WEDEL
University of Michigan

BARTON A. WEITZ
University of Florida

RUSSELL S. WINER
New York University

ANTHONY ZAHORIK
Burke Institute

VALARIE A ZEITHAML
University of North Carolina, Chapel Hill

Z. JOHN ZHANG
University of Pennsylvania

SHETH FOUNDATION/JOURNAL OF MARKETING AWARD

The recipients of the 2005 Sheth Foundation/Journal of Marketing Award, which is administered through the American Marketing Association Foundation, are **Donna L. Hoffman** and **Thomas P. Novak** for their article "Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations," which appeared in the July 1996 (Vol. 60, No. 3) issue of *Journal of Marketing.*

Nominations for the award were solicited from members of the *JM* Editorial Review Board, and a selection committee of prior editors of AMA journals made the final selection. In selecting the article, the selection committee observed that Hoffman and Novak's pioneering article has been widely cited in the academic literature and has had an important and lasting influence on marketing. As one of the letters of nomination noted, the article "offered a clear and testable theoretical perspective, drawn from high impact work in motivational psychology." Hoffman and Novak's stream of work has substantially and significantly influenced the field of marketing by stimulating the development of theory, empirical research, and improvement in the practice of marketing.

The Sheth Foundation/Journal of Marketing Award, which was established through the generosity of the Sheth Foundation, honors articles published in *Journal of Marketing* that have made long-term contributions to the *discipline of marketing*. The criteria for selection include the quality of the article's contribution to theory and practice, the article's originality, its technical competence, and its impact on the field of marketing. Articles eligible for the award must have been published at least five years but not more than ten years earlier.

The award, which carries both a plaque and a cash prize, was presented at the American Marketing Association's 2006 Winter Marketing Educators' Conference, held February 17–20, at the Tradewinds Island Resort in St. Petersburg, Fla.

- Donna L. Hoffman is Professor of Marketing and Co-Director and Co-Founder of eLab, a pioneering corporate-sponsored research center, at the Owen Graduate School of Management, Vanderbilt University.
- Thomas P. Novak is Professor of Marketing and Co-Director of eLab at the Owen Graduate School of Management, Vanderbilt University.

HAROLD H. MAYNARD AWARD

Deborah J. MacInnis and **Gustavo E. de Mello** have been selected as the recipients of the 2005 Harold H. Maynard Award for their article "The Concept of Hope and Its Relevance to Product Evaluation and Choice," which appeared in the January 2005 (Vol. 69, No. 1) issue of *Journal of Marketing*. The article was chosen by members of the *JM* Editorial Review Board for its significant contribution to *marketing theory and thought*.

The award, which carries both a plaque and a cash prize, was presented at the American Marketing Association's 2006 Winter Marketing Educators' Conference, held February 17–20, at the Tradewinds Island Resort in St. Petersburg, Fla.

•Deborah J. MacInnis is Department Chair and Professor of Marketing, Marshall School of Business, University of Southern California.

•Gustavo E. de Mello was a doctoral student in marketing, Marshall School of Business, University of Southern California.

MARKETING SCIENCE INSTITUTE/H. PAUL ROOT AWARD

Werner J. Reinartz, Jacquelyn S. Thomas, and **V. Kumar** have been selected as the recipients of the 2005 MSI/H. Paul Root Award for their article, "Balancing Acquisition and Retention Resources to Maximize Customer Profitability," which appeared in the January 2005 (Vol. 69, No. 1) issue of *Journal of Marketing*. The article was chosen by members of the *JM* Editorial Review Board for its significant contribution to the *advancement of the practice of marketing*.

The award, which carries both a plaque and a cash prize, was presented at the American Marketing Association's 2006 Winter Marketing Educators' Conference, held February 17–20, at the Tradewinds Island Resort in St. Petersburg, Fla.

A second outstanding article received numerous nominations for the award. For that reason, honorable mention is extended to Pierre Chandon, Vicki G. Morwitz, and Werner J. Reinartz for their article "Do Intentions Really Predict Behavior? Self-Generated Validity Effects in Survey Research," which appeared in the April 2005 (Vol. 69, No. 2) issue of *Journal of Marketing*.

•Werner J. Reinartz is Associate Professor of Marketing, INSEAD.

•Jacquelyn S. Thomas is Associate Professor of Integrated Marketing Communications, Northwestern University.

•V. Kumar is ING Chair Professor and Executive Director, ING Center for Financial Services, School of Business, University of Connecticut.

Raji Srinivasan, Gary L. Lilien, & Arvind Rangaswamy

The Emergence of Dominant Designs

In many product categories, technological evolution results in the emergence of a single product design that achieves market dominance. In this article, the authors examine two questions: Will a dominant design emerge in a new product category? and If it does, how long will it be before a dominant design emerges? Thus, the authors simultaneously model the probability of emergence of a dominant design and the time of that emergence, conditional on its emergence. The model incorporates the effects of several product-market characteristics on the probability and time of dominant design emergence, including appropriability of the rents associated with the product, network effects, size of the product's value net, the standards-setting process, radicalness of innovation, and research-and-development intensity. The authors use data for 63 office products and consumer durables to estimate a split-population hazard model for the probability and time of emergence of the dominant design. They find that a dominant design is more likely to emerge with weak appropriability, weak network effects, low product radicalness, and high research-and-development intensity. Dominant designs that emerge are likely to emerge sooner in product categories in which there is weak appropriability, there are a large number of firms in the value net, the standards are set by a de facto process, and there is low product radicalness The proposed model can be used to predict both the probability and the time of the emergence of a dominant design in a new product category.

In many new product categories, the market accepts a particular product's design architecture as one that defines the specifications for the entire product category; such a design is referred to as the dominant design (Abernathy and Utterback 1978; Utterback 1994). The emergence of a dominant design in a product category is an important event that significantly affects firms' strategies and performance. The late emergence of a dominant design can retard market growth because potential customers delay purchasing the product and wait for a dominant design to emerge (Utterback 1994). A dominant design also shapes future generations of products in the category, resulting in "an architectural franchise" for the firm with a dominant design and potentially locking out competitors (Schilling 1998). In this article, we examine the probability and time of emergence of a dominant design in a new product category.

There is considerable variation in the emergence of dominant designs across product categories. For example, the modern fax machine was introduced in 1960, but the dominant design, the GIII, did not emerge until 1983 (Baum, Korn, and Kotha 1995). Conversely, the dominant design for DVD players emerged just three years after product introduction in 1996. In other product categories, a dominant design has not emerged many years after product introduction, and it seems unlikely that a dominant design will ever materialize. For example, the camcorder category, introduced in 1984, supports multiple designs more than 20 years later. These designs include the VHS (home video system), VHS-C (home video system–compact), super VHS, super VHS-C, 8 mm, Hi-8, and digital 8 mm formats.

Dominant designs have been documented in diverse product categories, including VCRs, nuclear reactors, automatically controlled machine tools, and watches (Utterback 1994). The factors that influence the emergence of dominant designs have been identified with conceptual models (Lee et al. 1995; Smith 1997; Suaréz 2004), whereas empirical research has focused primarily on the *ex post* performance implications of dominant designs (Christensen, Suaréz, and Utterback 1998; Suaréz and Utterback 1995).

The study of dominant designs is highly relevant for marketing managers whose strategies must be informed by marketplace dynamics. Competition before the emergence of a dominant design is between designs, whereas after its emergence, competition is within the more circumscribed domain of a dominant design. A dominant design not only influences the strategies and performance of firms in the product category (e.g., DVD players) but also affects firms in ancillary categories (e.g., movies and video games for DVD players).

In this article, we examine the effects of product-market characteristics on the probability and time of the emergence of a dominant design. We address two questions: First, what

---

Raji Srinivasan is an assistant professor, Red McCombs School of Business, University of Texas at Austin (e-mail: Raji.Srinivasan@mccombs.utexas.edu). Gary L. Lilien is Distinguished Research Professor of Management Science (e-mail: GLilien@psu.edu), and Arvind Rangaswamy is Jonas H. Anchel Professor of Marketing (e-mail: arvindr@psu.edu), Smeal College of Business, Pennsylvania State University. The authors gratefully acknowledge the support of Pennsylvania State's Institute for the Study of Business Markets. The authors thank Reuben Raj, Barry Bayus, Murali Chandrashekaran, Rajesh Chandy, Peter Golder, Gerard Tellis, and Ken Krechmer for useful feedback on a previous version of the manuscript. The authors also thank the participants at the 2004 Marketing Science conference at Rotterdam and at Marketing Science Institute's Hot Thoughts on Innovation Conference for useful comments on the research and the anonymous *JM* reviewers for their insightful and constructive comments.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

product-market characteristics influence whether a dominant design will emerge in a product category? Second, conditional on the emergence of a dominant design, what product-market characteristics influence the time of that emergence?

Consistent with prior research (Christensen, Suaréz, and Utterback 1998), we take a market-facing view and define the time of the dominant design's emergence as when the market clearly favors one design over others. Unlike prior research, which has implicitly assumed that a dominant design will always emerge, our approach incorporates the possibility that a dominant design may sometimes never emerge.

We estimate a split-population hazard model (Schmidt and Witte 1989; Sinha and Chandrashekaran 1992) of the probability and time of emergence of a dominant design, which allows for two subpopulations of product categories: one in which a dominant design will emerge and one in which it will never emerge. We incorporate product-market characteristics as predictors of the emergence of dominant designs and calibrate the model using data on 63 office products and consumer durables. The data are right censored for 33 products for which a dominant design had not yet emerged by 2003.

The results show that a dominant design is more likely to emerge with weak appropriability of rents associated with the product, weak network effects, low product radicalness, and high research-and-development (R&D) intensity. Dominant designs that emerge do so sooner when there is weak appropriability, a large number of firms in the value net, de facto standards, and low product radicalness. The same factor may influence both the dominant design's probability and its time of emergence, but in different ways. The proposed model can be used as a predictive tool to estimate both the probability of emergence of a dominant design and the time of its emergence, given a set of product-market characteristics.

We organize the article as follows: In the next two sections, we provide an overview of dominant designs and present the proposed model and hypotheses. We then describe the data, the model estimation procedures, and the results. We conclude by discussing the article's contributions, summarizing its limitations, and identifying directions for further research.

Dominant Designs: An Overview

The concept of a dominant design emerged from studies on industrial innovation in the 1970s. In the early stages of market evolution, high market and technical uncertainty results in a diversity of product designs (Abernathy and Utterback 1978; Dosi 1982; Lee et al. 1995; Smith 1997; Utterback 1994). At some point, the market favors one product's design (i.e., the future dominant design) over the other designs. A dominant design may not necessarily be the one that embodies superior technical performance; sometimes, it is a satisficing design in terms of technical possibilities driven by the accommodation of commercial interests among suppliers, users, and competitors (Tushman and Rosenkopf 1992; Wade 1995).

Evolutionary processes of variation, selection, and retention characterize product category evolution (e.g., Anderson and Tushman 1990; Tushman and Murmann 1998). Technological breakthroughs create a rivalry between alternative designs, resulting in a period of design variation or ferment. Viewed in this evolutionary perspective, the emergence of a dominant design is the transition point between the periods of variation and selection.

Empirical studies document the emergence of dominant designs in diverse product categories, including typewriters, televisions, electronic calculators, automobiles (Utterback 1994), VCRs (Cusumano, Mylonadis, and Rosenbloom 1992), cochlear implants (Van de Ven and Garud 1994), fax machines (Baum, Korn, and Kotha 1995), cement, glass, and minicomputers (Anderson and Tushman 1990). The emergence of a dominant design affects firms' market shares (e.g., Anderson and Tushman 1990) and survival (e.g., Baum, Korn, and Kotha 1995; Christensen, Suaréz, and Utterback 1998; Suaréz and Utterback 1995; Tegarden, Hatfield, and Echols 1999).

However, dominant designs do not emerge in all product categories (e.g., camcorders, supercomputers, video game consoles). To our knowledge, prior research has not examined why dominant designs emerge in some product categories but not in others. Thus, we address the question in this article.

Definition of Dominant Designs

Over the years, the definition of dominant designs has evolved from being broad and possibly tautological (i.e., being defined by its effects) to being more specific. Table 1 summarizes alternative definitions of dominant designs in the literature. In this article, we seek a definition of a dominant design that is independent of its effects. Following the work of Christensen, Suaréz, and Utterback (1998), we define a dominant design as the specification (consisting of a single design feature or a complement of design features) that defines the product category's architecture. We measure the time to a dominant design's emergence as the time between the introduction of the first product in the category and the time an alternative in the category satisfies our definition.

Dominant Designs Versus Standards

Some prior research (e.g., Anderson and Tushman 1990; Besen and Farrell 1994; Katz and Shapiro 1986; Schilling 1998) has equated "dominant designs" with "standards," using the terms interchangeably. Indeed, such terms as "standards wars" have been used to denote the battle between designs that results in a dominant design (Shapiro and Varian 1999). In this article, we draw a distinction between dominant designs and standards.

Following the convention used in the engineering discipline and by numerous standards bodies, we use the term "standards" to denote the technical specifications for quality, reference, compatibility, adaptability, and connectivity that are required for the proper functioning of products, such as fax machines, computers, and railway tracks (Grindley 1995; Krechmer 2000). Standards are an

TABLE 1
Alternative Definitions of a Dominant Design in the Extant Literature

Source	Definition of a Dominant Design	Empirical Method to Identify a Dominant Design
Abernathy and Utterback (1978)	A dominant design is a single architecture that establishes dominance in a product category.	Conceptual paper.
Anderson and Tushman (1990)	A dominant design is a single architecture that establishes dominance in a product category.	A design is dominant if it acquires more than 50% market share of the product category and maintains it for four consecutive years.
Utterback (1994)	The dominant design in a product category is the one that wins the allegiance of the marketplace; it is the one that competitors and innovators must adhere to if they hope to command significant market following. A dominant design is a product in a product category that gains general acceptance as the standard on technical features that other market players must follow if they wish to acquire significant market share.	No details provided.
Suaréz and Utterback (1995)	The dominant design is a specific path along an industry's design hierarchy that establishes dominance among competing design paths.	Industry experts were used to classify dominant designs in typewriters, automobiles, televisions, picture tubes, transistors, and electronic calculators.
Christensen, Suaréz, and Utterback (1998)	A dominant design emerges in a product category when one product's design specifications (consisting of a single or a complement of design features) define the product category's architecture.	Industry experts were used to identify the emergence of the dominant design in the rigid disk drive industry based on the technical elements of the product category evolution over time.

inevitable requirement for these products because of interdependency among several components or among the products' users because these products must either connect with one another directly (e.g., fax machines) or indirectly (e.g., personal computers [PCs]) or require complementary goods to meet users' needs (e.g., DVD players). Thus, standards in a product category serve a functional purpose that is independent of market acceptance, whereas by definition, market acceptance is an integral aspect of a dominant design.

Two other aspects of dominant designs differentiate them from standards: First, a dominant design often emerges from competition among several designs. For example, in the home video recorder market, the VHS emerged as the dominant design as a result of competition between two video standards, Sony's Betamax and Japanese Victor Company's (JVC's) VHS. Second, dominant designs often consist of numerous standards. For example, the dominant design in the CD-player category, the Philips–Sony design, has more than two dozen codified standards.

To summarize, although there are product categories in which dominant designs are equivalent to the standard in the product category (e.g., DVD in digital home video players), there are also categories with multiple standards and no dominant design (e.g., for camcorders, there are VHS, 8 mm, mini-DVD, and so forth).

A Model of the Probability and Time of Emergence of a Dominant Design

Technological evolution has long been viewed as if it were a random process with a path-dependent trajectory that defies systematic modeling (e.g., Ames and Rosenberg 1977; Rosenberg 1994). However, this view has been challenged by studies that model a product category's evolution as a function of its product-market characteristics (Agarwal and Bayus 2002; Golder and Tellis 1997; Gupta, Jain, and Sawhney 1999). In the spirit of this latter work, we suggest that a product category's product-market characteristics affect its evolution, influencing both the probability and the time of the emergence of a dominant design.

Extending previous research, we argue that the following product-market characteristics affect whether and when a dominant design emerges in a product category: appropriability, network effects, the number of firms in its value net, the type of standards-setting process (de jure versus de facto), its radicalness, and its R&D intensity. We suggest that some product-market characteristics affect both the probability and the time of a dominant design's emergence, whereas others affect only the probability or the time of emergence. Next, we define these product-market characteristics and develop hypotheses of their effects on the probability and time of emergence of a dominant design.[1]

Appropriability

Appropriability refers to aspects of the product category that govern firms' abilities to capture innovation rents in the product category (Levin et al. 1987; Teece 1986). Levin and colleagues (1987) studied more than 100 industries and concluded that appropriability is a multifaceted construct that embodies the following six aspects: patents to prevent duplication, patents to secure royalty income, secrecy of the new product development effort, lead time for development of the product, learning-curve efficiency, and sales and service effort that underlies the innovation. Appropriability for a product category can be represented along a continuum that extends from a tight (closed) regime, in which firms can appropriate most or all innovation rents, to a weak (open) regime, in which firms can appropriate little or no rents.

Tight appropriability leads to disjointed competition and localized monopolies with several independent, disconnected market niches that reduce the selection pressures that are crucial for the emergence of a dominant design (Anderson and Tushman 1990; Frenken, Saviotti, and Trommetter 1999). With tight appropriability, product prices may also be higher because of limited industrywide learning effects, lowering the probability of emergence of a dominant design (Levin et al. 1987). In such a situation, developers of complementary goods may not have the knowledge or the incentives to develop complementary goods, reducing selection pressures and delaying the emergence of a dominant design.

The advantages associated with a weak appropriability regime have gained prominence with the emergence of the open-source movement in software development (Garud and Kumaraswamy 1993; Lecocq and Demil 2002; West 2003). Weak appropriability enables knowledge sharing among firms and the coevolution of networks of cooperating competitors (Brandenberger and Nalebuff 1996). With weak appropriability, all firms in the network, or the "value net" as we call it, can share in the innovation rents. Such rent-sharing reduces firms' incentives to innovate independently of the future dominant design, both increasing the probability of the emergence of a dominant design and hastening the time to emergence (Gallini 1985). For example, the emergence of the IBM PC as the dominant design in PCs resulted, in part, because of its open architecture with extensive public documentation (Khazam and Mowery 1994). Likewise, the emergence of the DVD format resulted, in part, because of the open format and extensive public documentation.

Integrating these arguments, we expect that weak appropriability in a product category increases selection pressures and positively affects both the probability of the emergence of a dominant design and the time of the emergence. Thus, we expect the events "emergence" (e) and "time of emergence" (t) to be conditionally dependent, given appropriability.

> H_{1e}: The weaker the appropriability of the product category, the more probable is the emergence of a dominant design.
>
> H_{1t}: The weaker the appropriability of the product category, the shorter is the time to the emergence of a dominant design.

Network Effects

Positive network effects exist when a customer's utility from a product increases as the number of customers who use the product (or compatible products) increases (Katz and Shapiro 1986). When the utility of a product to each user depends on the number of other users, the product exhibits direct network effects (e.g., fax machine, telephone). Indirect network effects arise when the link between consumer utility and the number of users in the network occurs through the increased availability of complementary products (e.g., movies on DVD for DVD players). An increase in the number of users of the focal hardware product increases the availability of complementary goods, which in turn increases the utility that customers derive from the focal product.

Some aspects of strong network effects suggest that dominant designs are more likely to emerge and emerge early. For example, a dominant design in a product category with strong network effects assures potential adopters that the dominant design's network will be the largest, resulting in its further adoption and ultimately establishing its market supremacy. Networked markets are typically "winner take all" so that a dominant design may have an early, large market share (Schilling 2002; Shapiro and Varian 1999). Such markets may tilt toward one design and do so rapidly because of selection pressures to choose a design early (e.g., Besen and Farrell 1994).

Other aspects of strong network effects suggest that dominant designs are less likely, and if they do emerge, they emerge later. Prospective customers of products with direct network effects may adopt a "wait-and-see" attitude, resulting in "excess inertia," which reduces the probability of emergence (Farrell and Saloner 1986; Goldenberg, Libai, and Muller 2002). Such excess inertia may also exist in product categories with indirect network effects in which the lack of complementary goods may cause a "chicken-and-egg" coordination problem that delays the emergence of a dominant design (Gupta, Jain, and Sawhney 1999). For example, in the HDTV market, Gupta, Jain, and Sawhney (1999) find that complementors' inactions delayed the emergence of a dominant design. Srinivasan, Lilien, and Rangaswamy (2004) report that the survival duration of a pioneer is negatively influenced by network effects (i.e., pioneers that introduce products with stronger network effects have shorter survival duration than pioneers that introduce products with weaker network effects), supporting the existence of excess inertia in product categories with network effects. Given the opposing effects of network effects, a priori, we do not hypothesize a directional effect of network effects on either the probability or the time of emergence of a dominant design; rather, we observe it empirically.

---

[1] The model also includes two control variables—the year of first product introduction and the presence of a recessionary environment during first product introduction—for which we do not develop hypotheses but present the results subsequently.

> H_{2e}: The stronger (weaker) the network effects in the product category, the less (more) probable is the emergence of a dominant design.

> H_{2t}: The stronger (weaker) the network effects in the product category, the longer (shorter) is the time to the emergence of a dominant design.

Value Net

A firm's linkages with its suppliers and producers of complementary products are important sources of relational rents for the firm and utility for its customers (Dyer and Singh 1998). The product's value net consists of suppliers and producers of complementary goods that deliver utility either directly or indirectly to the customer (Stabell and Fjeldstad 1998). Tushman and Rosenkopf (1992) and Wade (1995) suggest that the interdependencies within a value net, or what they call a "technological community," play a major role in the trajectory of technological evolution. Our interest is whether and how the size of the value net affects the emergence of a dominant design in a product category.

Because the value net typically consists of firms with different, frequently competing objectives, firms may initially prefer different designs, and thus we do not expect the value net of a product category to have the cohesion necessary to create selection pressures to affect the probability of the emergence of a dominant design; however, if the emergence of a dominant design seems inevitable, we expect the product's value net to affect the time to emergence, as we describe subsequently.

The more firms that are in the value net, the greater is the incentive for each firm to support a dominant design (if it emerges) because of higher resultant revenues (Amit and Zott 2001), thus reducing the time to the emergence of a dominant design. Consider the DVD player: Firms in the value net included the DVD player manufacturers (e.g., Matsushita, Sony, Philips), film studios (e.g., Paramount, Disney), film producers (e.g., Dreamworks, Pixar), movie rental chains (e.g., Blockbuster, Hollywood Video), and media manufacturers (e.g., Memorex, TDK). A dominant design permits firms in the value net to suspend further investments in product development, to compete within the dominant design, and to achieve supply-side and demand-side efficiencies (Utterback 1994). As a result, as the number of firms in the value net increases, the greater is the revenue generation in the value net, providing the financial and institutional pressure to establish a dominant design quickly (Wade 1995).

However, the more firms that are in the value net, the greater is the agency costs (e.g., transaction and coordination costs) for various firms in the value net, which could potentially delay a dominant design's emergence (Bergen, Dutta, and Walker 1992). Although we acknowledge this latter, opposing effect, we expect that the incentives for rent generation and sharing dominate the potential higher transaction costs.

> H_{3t}: The greater the number of firms that are in the value net of the product category, the shorter is the time to the emergence of a dominant design.

Note that according to our hypothesis, we expect the events of emergence and time of emergence to be conditionally independent, given the value net of the product category.

Standards-Setting Process

A review of the standards-setting process suggests that this process can be broadly classified into two types: de jure and de facto standards setting (Besen and Farrell 1994; David and Greenstein 1990; Farrell and Saloner 1988, 1992). In a de jure standards-setting process, standards are established through formal processes by standards organizations, which may include independent bodies, such as the International Telecommunications Union and the Institute of Electric and Electronic Engineers, or government bodies, such as the National Institute of Standards and Technology. In a de facto standards-setting process, standards for the product category are set by market forces. The operating system for PCs is an example of a product category that initially had several de facto standards (e.g., CP/M, MS-DOS). Because standards-setting processes, regardless of whether they are de jure or de facto, involve considerable uncertainty (e.g., compatibility, connectivity) among several firms with frequently competing interests, we do not expect the standards-setting process to affect the probability of the emergence of a dominant design, but we expect it to affect the time to the emergence of a dominant design (if it emerges).

De jure standards-setting processes, which typically include manufacturers, suppliers, and complementors, exclude end users and tend to support current or known technologies over emergent or new ones (Sirbu and Zwimpfer 1985). Because there are many member participants in the standards organization with potentially conflicting interests, it is more difficult to achieve consensus, delaying the emergence of a dominant design (David and Greenstein 1990). The fax machine standard was set through a de jure standards-setting process by several members of the Consultative Committee for International Telegraph and Telephone located in several different countries, and it required 23 years for the emergence of a dominant design.

In contrast, de facto standards-setting processes are characterized by aggressive market competition without intervention by standards bodies. Possible unilateral control of de facto standards by one firm, whose design may likely be the future dominant design, could also alleviate consumer uncertainty about a war of attrition between standards and shorten the time to the emergence of a dominant design (Shapiro and Varian 1999). Consistent with these arguments, Farrell and Saloner (1988) find that de facto standards-setting processes are faster than de jure processes.

> H_{4t}: Compared with a de jure standards-setting process, a de facto standards-setting process in a product category results in a shorter time to the emergence of a dominant design.

Radicalness

Radical products (e.g., instant photography, microwave ovens) involve new technologies that offer significant advances in both technology and consumer benefits (Chandy and Tellis 2000). Radical products, which offer a low initial performance-to-price ratio, face limited market acceptance (Christensen 1997; Utterback 1994). The evolution of radical products is characterized by high uncertainty, and numerous product variants are developed to identify viable technological trajectories and consumer preferences (Sorenson 2000). Not surprisingly, a radical product is refined by several firms in sequential iterations into a "market-ready" form (Basalla 1988). The limited market acceptance for rudimentary versions of the radical product may dampen selection pressures for a dominant design's emergence, thus making it less likely to emerge (Tushman and Murmann 1998). For example, the microwave oven, which Raytheon developed in 1946 and Tappan commercialized in 1955, was a radical product that faced poor initial market acceptance. Independent development efforts by several U.S. and Japanese firms over 20 years were necessary before the introduction of market-worthy designs.

The evolution of a radical product occurs over several years and involves the sequential development of several technology platforms (Rosenberg 1994). The development of variants, each with different marketplace support, leads to local, idiosyncratic market niches with independent product development efforts, thus delaying the emergence of a dominant design. Thus, we expect a negative effect of radicalness on the probability of the emergence of a dominant design and a positive effect on its time of emergence.

> H_{5e}. The greater the radicalness of the product category, the less probable is the emergence of a dominant design.
>
> H_{5t}: The greater the radicalness of the product category, the longer is the time to the emergence of a dominant design.

R&D Intensity

Product categories differ in their R&D intensity, that is, the depth and breadth of knowledge required to design and commercialize a product (John, Weiss, and Dutta 1999). In R&D-intensive product categories, widespread integration of interdisciplinary technologies across diverse firms may be necessary to achieve the technology fusion that is critical for product development (Iansiti and West 1997). Thus, a dominant design can provide a clear pathway in what may otherwise be a noisy and confusing technological environment.

Product categories that are R&D intensive are characterized by high levels of knowledge spillovers, not only across firms in the category but also across related categories (Tsai 2005). Knowledge spillovers in such product categories spur advances in multiple market niches, increasing product diversity, which in turn creates the variation that is crucial for the emergence of a dominant design (Wilson, Weiss, and John 1990). Moreover, product categories with higher R&D intensity also experience frequent technological changes, resulting in rivalries between alternative designs (Adam, Trajtenberg, and Fogarty 2000; Lotz 1998), which increases the selection pressures for the emergence of a dominant design. However, given the noisy and confusing technology environment in R&D-intensive product categories that persist even as the dominant design emerges, we do not expect R&D intensity to affect the time of emergence of the dominant design.

> H_{6e}: The greater the R&D intensity of the product category, the more likely is the emergence of a dominant design.

Table 2 summarizes the hypotheses for the probability and the time of the emergence of a dominant design.

Method

Split-Population Hazard Model

To allow for the possibility that a dominant design may never emerge, we use a split-population hazard model, which segments the population into two subpopulations, one that will experience the event (the emergence of a dominant design) and one that will never experience it. The model simultaneously estimates coefficients for both the probability that the event will ever occur and the event's time, conditional on its occurrence. The split-population approach has been typically applied to problems of criminal recidivism (e.g., Schmidt and Witte 1989). In marketing, Sinha and Chandrashekaran (1992) use a split-population model of the determinants of banks' technology adoption, allowing some banks to be eventual nonadopters.[2]

For each of the N product categories i, we observe the duration t_i for products for which a dominant design occurred in the interval $[0, \tau]$ and the vector X_i of covariates. Let C_i be an indicator that denotes whether a dominant design occurred in $(0, \tau)$ for the ith product category; $C_i = 1$ if the observation is complete, and $C_i = 0$ if it is censored. The data are characterized by Type I censoring (right censoring) so that each product category has a fixed censoring time $\tau > 0$; t_i is observed if $t_i \le \tau$; otherwise, $t_i > \tau$. Thus, $P(C_i = 1, t = t_i) = f(t_i)$ for the completed observations, and $P(C_i = 0) = P(t_i > \tau|X_i) = 1 - F(\tau) = S(\tau)$ for the censored observations. The log-likelihood function is

$$(1) \qquad LL = \sum_{i=1}^{N} \{C_i \ln[f(t_i)] + (1 - C_i)\ln[S(\tau)]\}.$$

Let $A = \{\ldots A_i, \ldots\}$ be an unobserved vector, where $A_i = 1$ if the ith product category witnesses the emergence of a dominant design and $A_i = 0$ if it will never witness a dominant design. We specify the probability of a dominant design's emergence as a function of the independent variables X_i: $\delta(X_i) = P(A_i = 1) = 1/[1 + \exp(\alpha X_i)]$, and $P(A_i = 0) = 1 - \delta(X_i)$. For product categories in which the dominant design emerges ($C_i = 1$), we also observe the time of its emergence, whereas when a dominant design does not emerge ($C_i = 0$), either a dominant design will never occur (i.e., $A_i = 0$) or a dominant design will eventually emerge, but beyond the observation period, in which case $t > \tau$. Thus, if a dominant design emerges during the observation period,

---

[2] For details on the statistics for the split-population model, see Schmidt and Witte (1989); for an application in a marketing context, see Sinha and Chandrashekaran (1992).

TABLE 2
Summary of Hypotheses and Results

Variable	Probability of Emergence of a Dominant Design	Time of Emergence of a Dominant Design
Appropriability	H_{1e}: The weaker the appropriability of the product category, the more probable is the emergence of a dominant design. (Supported)	H_{1t}: The weaker the appropriability of the product category, the shorter is the time to the emergence of a dominant design. (Supported)
Network effects	H_{2e}: The stronger the network effect in the product category, the less (more) probable is the emergence of a dominant design. (Less likely supported)	H_{2t}: The stronger the network effect in the product category, the longer (shorter) is the time to the emergence of a dominant design. (Not supported)
Value net	—	H_{3t}: The greater the number of firms in the value net of the product category, the shorter is the time to the emergence of a dominant design. (Supported)
Standards-setting process	—	H_{4t}: Compared with a de jure standards-setting process, a de facto standards-setting process in a product category results in a shorter time to the emergence of a dominant design. (Supported)
Radicalness of innovation	H_{5e}: The greater the radicalness of the product category, the less probable is the emergence of a dominant design. (Supported)	H_{5t}: The greater the radicalness of the product category, the longer is the time to the emergence of a dominant design. (Supported)
R&D intensity	H_{6e}: The greater the R&D intensity of the product category, the more probable is the emergence of a dominant design. (Supported)	—

$$(2)\qquad P(C_i = 1, t = t_i) = P(A_i = 1)f(t_i|X_i, A_i = 1) = \delta(X_i)f(t_i|X_i, A_i = 1),$$

and for censored observations,

$$(3)\qquad P(C_i = 0) = 1 - \delta(X_i) + \delta(X_i)[1 - F(\tau|X_i, A_i = 1)] = 1 - \delta(X_i) + \delta(X_i)S(\tau).$$

The log-likelihood function in Equation 1, when expressed as a function of the covariates X_i, is

$$(4)\qquad LL = \sum_{i=1}^{N}\{C_i(\ln[\delta(X_i)f(t_i|X_i, A_i = 1)] + (1 - C_i)\ln[1 - \delta(X_i) + \delta(X_i)S(\tau)]\}.$$

To estimate the effects of the covariates on the probability and the time of the emergence of a dominant design, we used the accelerated failure time specification, exploring alternative distribution functions $f(\cdot)$ for the base hazard. We implement the model within LIMDEP 8.0 (Greene 2002).

Data

We identified two product categories, office products and consumer durables, which have been extensively studied in prior research on market evolution (e.g., Agarwal and Bayus 2002; Golder and Tellis 1997). We used the historical method to collect data on the products (Golder 2000). We obtained information on the product categories from articles published in scholarly journals, books, and online business databases. We were able to collect complete information for 63 product categories (see Appendix A). All the data are from public sources.

Measures

Emergence and time of dominant designs. As we discussed previously, before the emergence of a dominant design, there are several technological changes in the product category. However, after a dominant design emerges, competition moves from competition between alternative product designs to competition within the dominant design. Guided by the distinctive characteristics surrounding the emergence of a dominant design, strategy researchers (Christensen, Suarẽz, and Utterback 1998) have retrospectively identified a dominant design on the basis of when one design achieves market acceptance over other competing designs. We use this method to identify the emergence of dominant designs.

Three graduate students used archival sources (e.g., *Consumer Reports*, trade magazines, computing magazines) to develop histories of the evolution of product categories (for brief histories of two such categories, see Appendix B).

We then used two other graduate students to identify whether and when a dominant design emerged.

Following the convention for survival models (Lawless 2003), we used two variables to code the emergence of a dominant design. The first denotes whether a dominant design emerged in the product category. If a dominant design emerged, the variable was coded as 1, but if a single design specification was not favored by the market (e.g., video game consoles, camcorders), the variable was coded as 0. We observed dominant designs in 30 product categories. For the 33 product categories in which a dominant design had not yet emerged in 2003, the data were right censored. We define the second variable as the number of years between the introduction of the product category and the emergence of a dominant design. There were discrepancies between the two raters for 3 of the 30 product categories for the time measure, which we resolved using inputs from industry experts.

Given its centrality in our study, we used two checks of the measure of a dominant design. First, we used 11 industry experts to verify the year of the dominant design's emergence obtained from archival sources. The experts included five editors of trade journals (e.g., *Disk Drive Report*, *Computer Reseller News*, *Future Image Report*), two representatives of trade organizations (Consumer Electronics Association, DVD Forum), and four executives (e.g., Gartner Group, IDC). In most cases (26 of the 30 product categories in which a dominant design emerged), there was unanimous agreement between student raters and industry experts. In four product categories (fax machine, floppy drive, mainframe computer, and microprocessor chip), the dates for the dominant design's emergence differed by, at most, two years, which we reconciled by consulting with the industry experts.

Second, following the work of Anderson and Tushman (1990), we validated the measure of a dominant design with an objective measure of its market share performance. We obtained archival measures of market share for the dominant design for 24 of the 30 products in which a dominant design emerged. Using Anderson and Tushman's method, we coded the year in which the dominant design's market share exceeded 50% for four consecutive years. For these 24 products, the dominant design's time measure correlates highly ($\rho = .86$) with the time to the 50% market share threshold, reassuring us of the reliability of our measure.

Because there are no objective measures for appropriability, network effects, value net, and radicalness, we used raters to develop measures for these product characteristics. When feasible, we validated these measures with data from secondary sources.

Appropriability. Levin and colleagues (1987) conceptualize a product's appropriability as protection of R&D advantages for firms in the product category, covering the following six aspects: (1) patents (or copyrights) to prevent duplication, (2) patents (or copyrights) to secure royalties, (3) secrecy of product development effort, (4) lead time for product development, (5) speed in moving down the learning curve in product development, and (6) sales and service support. Because the sixth aspect is not related to innovation, we do not include it in our measure of appropriability.

We used two groups of expert raters to measure appropriability: academic experts (six professors from three business schools who are recognized experts on organizational innovation) and industry experts (two sets). Following the work of Levin and colleagues (1987), we asked the experts to rate (on a seven-point scale) the products on appropriability. We averaged the ratings separately for the two groups of raters. The higher a product category's rating on the scale, the weaker is its appropriability.

Network effects. Following the work of Srinivasan, Lilien, and Rangaswamy (2004), we used raters to measure a product's network effects as a continuous, two-dimensional variable, representing both direct and indirect effects. Again, we used two groups of raters (different from the raters used for the other measures): academic experts (four professors at two business schools who are recognized experts on organizational innovation) and MBA students (25 students with an average of three years' work experience). We asked the raters to rate direct and indirect network effects separately for each product category on a seven-point scale, ranging from 1 ("no effects") to 7 ("very strong effects"). We computed network effects for each product from each rater by adding the scores for direct and indirect network effects and averaging separately for each group of raters.

Radicalness. We used Chandy and Tellis's (2000) radicalness measure, which has two dimensions: whether a new product incorporates a substantially different core technology (technology radicalness) and whether a new product provides substantially higher customer benefits relative to existing products (benefits radicalness). Both dimensions were assessed on nine-point scales. Six academics, all experts in organizational innovation, and ten graduate engineering students provided the radicalness ratings. We provided each rater with a brief history of each product category. We obtained the measure of radicalness for each product by adding each rater's ratings for technology and benefits radicalness and averaging the ratings separately for the two groups of raters.

Number of firms in the value net. Three graduate students, who prepared detailed case histories of the 63 products, also gathered details on the value net of the product category (Amit and Zott 2001). Two other graduate students (different from those who identified the dominant design) counted the number of firms in the value net. There was complete agreement between the two raters for 57 products (90% of the cases). The counts for the remaining 6 products were resolved by a discussion between the raters and one of the authors.

Standards-setting process. The two graduate student raters (those who counted the number of firms in the value net) also independently identified whether the standards-setting process was de facto (coded as 1) or de jure (coded as 0). There was unanimous agreement between the raters on the standards-setting process, which was reconfirmed with three academic experts.

R&D intensity. Consistent with prior research (Agarwal and Bayus 2002), we measured R&D intensity by the average R&D expenditure as a percentage of sales for each product category at the three-digit Standard Industrial Classification level.[3]

Recession at the time of product introduction. We measured recessionary environment using a categorical variable based on whether there was a recession in the U.S. economy in the year of first product introduction. We defined recession as two consecutive quarters of decline in gross domestic product.

Ratings from multiple raters. For the three explanatory variables of appropriability, network effects, and radicalness, multiple raters rated all the products. There are two ways that ratings from multiple raters may bias the estimates of the model. First, there is the potential for differences in the reliability of ratings across the raters (Van Bruggen, Lilien, and Kacker 2002). To account for this bias, we computed a weighted mean of the raters' ratings; the weight assigned to a rater was the reciprocal of the rater's absolute distance from the unweighted mean rating compared with other raters, as Van Bruggen, Lilien, and Kacker (2002) propose. Second, there may be a measurement error bias because the rating instrument may have induced each rater to provide a rating that differs from its "true" value.

We report results using the rater-bias adjusted ratings from academic experts, which accounts for differences in reliability of the ratings from multiple raters. As we report subsequently, we test for robustness of these estimates using ratings obtained from the other raters. In addition, we also implement a simulation procedure that examines the robustness of the parameter estimates to potential measurement errors.

Results

Descriptive Statistics

In 33 (52%) of the 63 product categories, no dominant design emerged by 2003. The average time to the dominant design's emergence for the 30 product categories in which a dominant design emerged was 6.50 years. Table 3 contains the descriptive statistics and correlation matrix of the variables. Correlations are within acceptable limits (highest = .56, between network effects and the size of the value net), and the variance inflation factors were lower than ten (average = 1.48; maximum = 1.94), suggesting that multicollinearity is not a threat to the validity of the findings (Belsley, Kuh, and Welsch 1980).

Model Selection and Results

We estimated the model in Equation 5 using three distribution functions (lognormal, log-logistic, and Weibull) that accommodate a changing hazard rate. Although the general pattern of results is similar across the models, the model estimated with the log-logistic function fits best. We report the results for that model in Table 4 (Column 1). The parameters of the log-logistic function ($\lambda = .23$; $\sigma = .31$) suggest that the hazard function is mound shaped, increasing for the first six years and then decreasing. Although dominant designs emerged in only 30 of the 63 product categories (48%) during the period of observation, the hazard model suggests that, on average, there is a 60% chance that a dominant design would eventually emerge in the office products and consumer durables categories (see last row of Table 4).

We first discuss the effects of the control variables. The year of product introduction has no effect on the probability of emergence (b = 5.63, not significant [n.s.]) or on the time of emergence of a dominant design (b = 4.30, n.s.), whereas the presence of a recessionary environment (b = .70, $p < .01$) delays the emergence of a dominant design. In general, the results support the hypotheses. With respect to the prob-

TABLE 3
Descriptive Statistics and Correlation Matrix of Key Variables (n = 63)

Variable	Range of Variable	M (SD)	2	3	4	5	6	7
1. Time to the emergence of a dominant design (n = 30)	2–20	6.50 (4.96)	—[a]					
2. Appropriability	1–7	3.20 (1.94)	1.00					
3. Network effects	2–14	8.52 (3.12)	–.46**	1.00				
4. Number of firms in the value net	1–5	2.21 (1.10)	–.18	.56***	1.00			
5. Standards-setting process	0–1	54–de facto 85%	–.04	–.24*	–.17	1.00		
6. Radicalness	2–18	11.89 (2.11)	.11	–.04	–.09	–.11	1.00	
7. R&D intensity	1.67–9.00	7.40 (2.76)	–.03	–.17	–.05	.50***	.18	1.00

*$p < .10$.
**$p < .05$.
***$p < .01$
[a]Correlations with the left-censored duration variable are not meaningful, and therefore we do not report them.

---

[3] We obtained data for the R&D intensity from the Web (National Science Foundation, Division of Science Resources Studies 1999).

TABLE 4
Split-Population Model for Emergence of Dominant Design

Variable	Hypothesized Model	Single-Population Duration Model	Cutoff Year = 2002	Cutoff Year = 1998	Bootstrapping
Probability of Emergence					
Constant	−32.32 (733.18)		−31.20 (748.92)	−33.65 (788.46)	−31.31 (743.08)
Appropriability (H_{1e})	1.51 (.57)***	—	1.52 (.59)***	1.42 (.69)**	1.50 (.56)***
Network effects (H_{2e})	−.60 (.28)**	—	−.63 (.24)**	−.61 (.29)**	−.59 (.30)**
Radicalness (H_{5e})	−1.47 (.63)**	—	−1.47 (.62)**	−1.48 (.62)**	−1.47 (.63)**
R&D intensity (H_{6e})	.81 (.48)*	—	.82 (.50)*	.82 (.59)	.82 (.50)*
Year of product introduction	5.63 (96.15)	—	5.68 (97.30)	5.76 (99.30)	5.61 (97.84)
Time of Emergence					
Constant	−33.13 (151.75)	40.51 (287.00)	−30.20 (157.96)	−28.45 (162.02)	−32.59 (150.08)
Appropriability (H_{1t})	−.21 (.10)**	.29 (.14)**	−.21 (.09)**	−.22 (.10)**	−.21 (.09)**
Network effects (H_{2t})	.08 (.08)	−.14 (.11)	.08 (.08)	.08 (.08)	.08 (.08)
Number of firms in the value net (H_{3t})	−.35 (.12)***	−.20 (.20)	−.32 (.14)**	−.32 (.15)**	−.34 (.12)***
Standards-setting process (H_{4t})	−.88 (.24)***	−.03 (.56)	−.89 (.23)***	−.88 (.25)***	−.88 (.25)***
Radicalness (H_{5t})	.28 (.07)***	−.08 (.11)	.29 (.07)***	.28 (.08)***	.28 (.05)***
Year of product introduction	4.30 (19.89)	4.74 (37.78)	4.21 (20.317)	4.33 (22.65)	4.20 (19.46)
Recession during product introduction	.70 (.22)***	.12 (.43)	.69 (.21)***	.69 (.22)**	.70 (.21)***
σ of log-logistic distribution	.31	.74	.29	.29	.30
Log-likelihood	−44.15	−68.72	−41.82	−39.46	−44.04
Average predicted probability of emergence (δ)	.60	—	.60	.60	.61

*$p < .10$.
**$p < .05$.
***$p < .01$.
Notes: We show the parameter estimates and standard errors in parentheses.

ability of the emergence of a dominant design in a product category, as we hypothesized in H_{1e}, we find that appropriability has a positive effect (b = 1.51, $p < .01$), suggesting that the weaker the appropriability, the more likely is the emergence of a dominant design. As we hypothesized in H_{6e}, the greater the R&D intensity, the more likely is the emergence of a dominant design, though we obtain only marginal significance for this result (b = .81, $p < .10$). Conversely, as we hypothesized in H_{2e} and H_{5e}, respectively, network effects (b = –.60, $p < .05$) and radicalness (b = –1.47, $p < .05$) negatively affect the probability of emergence of a dominant design.

With respect to the time of emergence, as we hypothesized in H_{1t}, H_{3t}, and H_{4t}, respectively, weak appropriability (b = –.21, $p < .05$), the number of firms in the value net of the product (b = –.35, $p < .01$), and a de facto standards-setting process (b = –.88, $p < .01$) hasten the emergence of a dominant design. As we hypothesized in H_{5t}, radicalness (b = .28, $p < .01$) delays the emergence of a dominant design. We find no support for H_{2t} for the effect of network effects on the time of the emergence of a dominant design (b = .08, n.s.).

Next, we performed two model comparisons to examine the explanatory power of the split-population hazard model. First, we compared the proposed split-population hazard model with all explanatory variables for both the probability and the time of emergence with a baseline model that excluded the explanatory variables. A likelihood ratio test of the difference between the model with no explanatory variables and the proposed model ($\chi^2 = 51.92$, d.f. = 12, $p < .01$) indicated that the proposed split hazard log-logistic model outperformed this baseline model. Second, we compared the proposed model with a single-population model (i.e., incorporating only time; see Column 2 of Table 4). The proposed complete model outperformed the single-population model ($\chi^2 = 49.14$, d.f. = 5, $p < .01$). In addition, the parameter estimates change sign, and only appropriability has a significant, though opposite, effect in the single-population model. Such changes in sign may be expected when a single-population model is a misspecification of an underlying split population (Sinha and Chandrashekaran 1992).

Robustness Tests

Model specification. We performed additional analyses to examine the robustness of the model specification. First, we examined the effect of stepwise addition of variables to our models (though we do not report the results here). For both the probability and the time submodels, the coefficients are consistent in terms of their signs in the stepwise models with those we report in Table 4, reconfirming that multicollinearity is not a threat to the validity of our findings. Second, we tested for the effects of two additional control variables (both dummy coded): whether the candidate designs in the product category had a dominant firm sponsor, which we measured by sales revenue, at the time of product introduction and whether the candidate designs in the product category were backward compatible with a previous generation product, as judged by two graduate student raters. The analysis indicated that these two variables affected neither the probability nor the time of the emergence of a dominant design, supporting the robustness of the model specification.

Censoring date. As we discussed previously, the censoring mechanism here is Type I, for which the maximum likelihood estimation of the hazard function provides consistent estimates (see Lawless 2003). To explore whether the results are sensitive to sampling variations due to the use of 2003 as the censoring date, we reestimated the model with alternative censoring dates: 2002, 2001, 2000, 1999, and 1998. We obtained consistent results, though, as may be expected, the significance levels of the parameter estimates varied across the different censoring dates. We report the results for the censoring dates of 2002 and 1998 in Columns 3 and 4 of Table 4.

Sampling variations. Although the sample size compares well with prior research, it is small; therefore, we performed a bootstrap analysis to examine the sensitivity of the results to sampling variations. The bootstrapping analysis (Column 5 of Table 4) indicated that the results were robust to sampling variations.

Use of multiple raters. To cross-validate the estimates obtained from the ratings of appropriability, network effects, and radicalness provided by the academic raters, we reestimated the model using the rater-bias adjusted ratings obtained from students and industry experts. These estimates were consistent with those we report in Table 4.

Effect of measurement errors. To assess the effects of measurement errors in ratings, we conducted a simulation (with 30 randomly generated samples) by assuming that the average ratings provided by the raters across different products are normally distributed in the population with mean and variance equal to the overall mean and variance (across all raters), respectively (i.e., $\bar{X} \sim N[\mu_{\bar{X}}, \sigma_{\bar{X}}]$), for each variable measured with error.[4] For the simulations, we set $\mu_{\bar{X}}$ equal to its observed mean in the sample, and $\sigma^2_{\bar{X}} = s^2_{\bar{X}}/(k - 1)$, where $s^2_{\bar{X}}$ is the observed variance in the sample and k is the number of raters. For each variable and for each rater, we drew a random number from $N(0, \sigma_{\bar{X}})$ and added that value to the rating of each product provided by that rater. We then computed the combined ratings across the raters as Van Bruggen, Lilien, and Kacker (2002) propose. We reestimated the split hazard model, once for each of the 30 random data sets generated, to obtain the mean and standard deviation of the estimates across the 30 simulations. The reestimated models provided results that were consistent with those we report in Table 4 (further details of the simulation procedure are available on request).

Direct and indirect network effects. We also separately examined the impact of direct and indirect network effects on the probability and time of emergence of a dominant design (though we do not report it here). The hypothesized model (Bayesian information criterion value = 143.49) outperformed the disaggregated network effects model (Bayesian information criterion value = 151.83). In the disaggregated model, direct network effects have a weak effect ($p < .01$) on the probability and time of emergence of a dominant design, whereas indirect network effects have no effect on either. We conjecture that our inability to account separately for the direct and indirect network effects on the emergence of dominant design may be because both direct and indirect network effects coexist ($\rho = .25$, $p < .01$) and jointly determine the dynamics of dominant design emergence.

---

[4] We pursued this approach because there is currently no available theory for the assessment of measurement error biases in continuous-time accelerated failure time models with multiple constructs, multiple populations, and right censoring (Augustin 2005)

Predictive Validity

Emergence of dominant design. Because the data are right censored at 2003, we may not observe the emergence of a dominant design in a product category during the observation period, because (1) a dominant design may never emerge in that category or (2) a dominant design may emerge after 2003. We selected a subset of product categories introduced in or before 1984 (n = 42) and used the model to predict the probability of emergence of the dominant design. The cutoff date of 1984 provides 20 years of observations after product introduction, or three standard deviations beyond the average time for dominant design emergence that we observed in our data. Thus, we can be reasonably certain that a dominant design is unlikely to emerge in those product categories after 2003.

The split-population log-logistic hazard model predicts (1) the emergence of a dominant design for 14 categories, 12 of which had witnessed the emergence of a dominant design by 2003, and (2) no emergence of a dominant design for 28 categories, 18 of which had not witnessed the emergence of a dominant design, resulting in a false-positive rate of 14% and a false-negative rate of 36%. By means of the criterion that Morrison (1969) advocates for evaluating the predictive performance of discriminant analysis models, the overall proportion of correct classification is .71, which compares favorably with the proportional chance C_{pro} of .50; $C_{pro} = [\alpha]^2 + [1 - \alpha]^2$, where $[\alpha]$ is the proportion of categories for which a dominant design has actually emerged and $[1 - \alpha]$ is the proportion for which a dominant design has not emerged. A chi-square test of misclassification indicates that the proposed split hazard model predicts dominant design emergence well ($\chi^2 = 12.92$, d.f. = 1, $p < .01$). We obtained similar results in the predictions of the emergence of a dominant design when we used alternative cutoff years of 1981 ($\chi^2 = 13.17$, d.f. = 1, $p < .01$) and 1987 ($\chi^2 = 10.32$, d.f. = 1, $p < .01$).

Time of emergence of a dominant design. To evaluate the model's ability to predict the time of emergence of a dominant design, we applied jackknifing, using data for the 30 products for which there was a dominant design. We held out one target product, reestimated the model on the other 29 products, and then used the estimated parameters to predict the emergence of a dominant design for the remaining products. We computed the mean absolute deviation (MAD), defined as $(1/n)\Sigma_{i=1}^{n}|D_{ia} - D_{ip}|$, where D_{ia}, D_{ip}, and n denote the actual time to the emergence of a dominant design for product i, the predicted time of the emergence of a dominant design for product i, and the number of products (n = 30), respectively.

The MAD using the predicted model is 2.49 years compared with a MAD without the model of 3.43 years, an improvement of 27%. A post hoc examination of absolute deviations indicates that elimination of two products (color television and audiocassette player) reduces the MAD for the predicted model to 2.20 years compared with a MAD without the model of 3.40 years, an improvement of 35%. Our small sample size precludes the determination of why these two products have such high leverage on model performance. Overall, the two predictive validity tests demonstrate that the split-population hazard model predicts both the probability and the time of the emergence of a dominant design well.

Discussion

We investigated the emergence and the time of emergence of a dominant design in a product category as a function of its product-market characteristics, allowing for the possibility that a dominant design may never emerge in some product categories. The model specifies the effects of product-market characteristics both on the probability of the emergence of a dominant design in a product category and on the time of the emergence of a dominant design, conditional on its emergence.

Theoretical Contributions

Our findings contribute to the literature in marketing that explores product and market evolution. The findings suggest that a split-population model that allows for the non-emergence of a dominant design in some product categories offers a better way to model the factors that influence the emergence of dominant designs than a single-population hazard model. The split-population model can be extended to other product and market evolution processes, such as product takeoff, to account for the possibility that some products may never take off (e.g., Sony minidisk, quadraphonic sound).

We find that dominant designs may never emerge in some product categories, a result that runs counter to an implicit assumption in the existing literature that a dominant design will eventually emerge in all categories. A product category's characteristics differentially influence the probability and time of emergence of a dominant design. The facilitating role of weak appropriability on both the probability and the time of emergence of a dominant design suggests that appropriability has opposing effects on firms' quests for securing rents from innovations (Teece 1986). On the one hand, weak appropriability results in a loss of rents because other firms can imitate the innovator's design. On the other hand, weak appropriability is associated with a greater probability of emergence and the earlier emergence of a dominant design, increasing rents for all firms in the market. The negative effect of network effects on the probability of the emergence of a dominant design reinforces recent empirical research (e.g., Goldenberg, Libai, and

Muller 2002; Srinivasan, Lilien, and Rangaswamy 2004) that suggests that for office products and consumer durables, the excess inertia of network effects appears to outweigh its installed base effects.

The negative effect of the number of firms in the value net on the time to the emergence of a dominant design suggests that if there are more firms appropriating rents in a product category, a dominant design is likely to emerge sooner. These findings are somewhat counter to other evidence on the adverse effects of coordination costs of interorganizational relationships. For the product categories in this study, the added rents from the value net appear to outweigh the added transaction costs, thus hastening the emergence of a dominant design.

Likewise, the results indicate that a dominant design is likely to emerge sooner in product categories in which the standards-setting process is de facto (versus de jure). This is consistent with prior theoretical work in economics (e.g., Farrell and Saloner 1988) and points to the efficiency of market mechanisms in forcing a quicker emergence of a dominant design.

Dominant designs are less likely to occur and, if they occur, take longer to emerge for more radical product categories. However, the results for R&D intensity suggest that dominant designs are more likely to occur in R&D-intensive categories; the higher investments in the product categories perhaps provide the necessary variation in designs, which in turn create pressures for the selection of a dominant design.

Managerial Contributions

The study's findings have some clear implications for managers. First, the results suggest that managers should realize that a dominant design may never emerge in many product categories, especially those with tight appropriability, strong network effects, and high product radicalness. Some of these characteristics are strategic decision variables. For example, weak appropriability significantly influences both the probability and the time of emergence of a dominant design. In addition, market-based de facto standards-setting processes (versus de jure standards-setting processes) appear to be efficient for the timely emergence of dominant designs. Managers who wish to increase the probability of their product becoming a dominant design and/or speed up the emergence of dominant designs might consider strategies to weaken appropriability and also encourage de facto standards-setting processes. For example, IBM and Intel's decision to share know-how for blade servers may be an attempt to hasten the emergence of a dominant design using a de facto standards-setting process (*The Wall Street Journal* 2004).

Second, the model can be used to predict the probability and time of emergence of dominant designs. These predictions can be useful to managers in the *ex ante* planning of marketing strategies. For example, of the products in our data set that have not yet had dominant designs emerge, the product category with the highest probability is satellite radio (p = .86), and the product category with the least likely probability is instant messenger (p = .01). It will be interesting to observe whether these predictions turn out to be accurate.

Limitations and Further Research

This study has some limitations that present opportunities for further research. We primarily focused on the time to the emergence of the initial dominant design in a new product category. Some researchers (e.g., Anderson and Tushman 1990; Tushman and Anderson 1986) have conceptualized the emergence of dominant designs as a sequence of technological discontinuities that result in several dominant designs. Indeed, in some product categories (e.g., operating systems for PCs, spreadsheet software), the initial dominant design was replaced by a subsequent design. To extend the findings from this study, further research could address the emergence of multiple, sequential dominant designs, perhaps through a repeated-events, split-population hazard modeling approach.

Given our focus on the development of a cross-product category model in which objective measures of some variables were not available, we used retrospective measures for some constructs, obtaining them from knowledgeable raters. With improved record keeping, future researchers could use measures that are not subject to the possible hindsight biases and measurement errors associated with these measures. Alternatively, researchers could consider focusing on a specific product category (e.g., Shankar and Bayus 2003), which may enable the use of objective measures within the product category.

The emergence of a dominant design may also be influenced by the strategic characteristics and actions of firms, including their collateral assets (Teece 1986), product portfolios, and strategic maneuvering (Cusumano, Mylonadis, and Rosenbloom 1992). Further research could extend our study by incorporating such firm-level factors.

We hope that we have contributed in some distinct way to the theory base that helps explain the emergence of dominant designs. We also hope that our approach and modeling framework are helpful for managers who need to devise strategies that reduce the uncertainty about the dynamics of evolution of markets for technology products.

APPENDIX A
Product Categories (n = 63)

Product Category	Year of Product Introduction	Time to the Emergence of a Dominant Design	Specifications of a Dominant Design
1. 3.5″ floppy drive	1979	1984	Sony's design
2. AM stereo	1982	1986	Motorola's C-Quam System
3. Answering machine	1962	—	—
4. Antivirus software	1982	—	—
5. Audiocassette player	1962	1969	Philips's design
6. Automatic teller machine	1967	—	—
7. Cable modem	1995	1998	DOCSIS specifications
8. CAD software	1982	—	—
9. Camcorder	1984	—	—
10. Camera phone	2000	—	—
11. CD player	1982	1985	Philips–Sony's design
12. CD-ROM drive	1983	1986	Sony's design
13. Cellular phone	1979	—	—
14. Color television	1951	1957	National Television System Committee
15. Cordless telephone	1975	—	—
16. Database software	1981	1986	Server Query Language (SQL)
17. Desktop publishing software	1984	1987	Adobe Pagemaker
18. Dial-up modem 56 K	1979	1998	56KFlex
19. Digital camera	1991	—	—
20. Dot matrix printer	1964	1968	ESC/P from Epson
21. DSL modem	1996	1999	G.Lite
22. DVD player	1996	1999	DVD design from Zenith
23. Electric toothbrush	1960	—	—
24. Fax machine	1960	1983	GIII
25. File-zipping software	1986	—	—
26. Flash memory	1988	—	—
27. Flatbed scanners	1978	—	
28. Food processor	1972	—	—
29. Graphics software	1990	1996	Adobe Photoshop
30. HDTV	1987	1993	Standard Definition Television
31. Home microwave oven	1955	—	—
32. Inkjet printer	1984	—	—
33. Instant messenger	1998	—	—
34. Instant photography	1948	1955	Polaroid
35. Internet service providers	1980	—	—
36. Internet telephony	1997	—	—
37. Laser printer	1984	—	—
38. Mainframe computer	1946	1964	IBM 360
39. Microprocessor chip	1971	1979	Intel 4004
40. Mini audio disc	1992	—	—
41. Notebook computer	1980	—	—
42. OCR software	1974	—	—
43. Operating system for PCs	1977	1984	MS DOS
44. Pager	1974	—	—
45. PC	1975	1983	IBM PC
46. Personal digital assistant	1993	—	—
47. Personal finance software	1983	1987	Intuit
48. Photocopiers	1950	1959	Xerox 914
49. Pocket calculator	1971	—	—
50. Portable file document software	1993	1999	Adobe PDF
51. Presentation graphic software	1986	1991	Harvard Presentation Graphics
52. Projection television	1973	—	—
53. Satellite radio	1998	—	—
54. Single-use camera	1986	—	—
55. Spreadsheet software	1979	1984	Lotus 1-2-3
56. Super audio CDs	1999	—	—
57. Video game console	1971	—	—
58. Home video recorders	1975	1978	JVC VHS
59. Web camera	1991	—	—
60. Web server software	1995	—	—

APPENDIX A
Continued

Product Category	Year of Product Introduction	Time to the Emergence of a Dominant Design	Specifications of a Dominant Design
61. Word-processing software	1979	1983	Wordstar
62. Work station	1980	1986	Sun's Unix
63. Zip drives	1995	1997	Iomega

Appendix B
Brief Histories of Two Product Categories

Camcorders (Dominant Design Did Not Emerge)

In January 1982, several leading consumer electronics firms announced a universal standard for camcorders, the 8 mm video format. In March 1983, 127 video-related companies agreed on standards to ensure compatibility of all 8 mm camcorders. Ten months later, Eastman Kodak Company introduced the world's first 8 mm camcorder in January 1984, followed by Sony in the fall of 1984.

In January 1985, several Japanese firms, including Canon, Fuji, Pioneer Electronic, Kyocera, Aiwa, Sanyo, and Ricoh, licensed Sony's 8 mm camcorder technology. In February 1985, JVC launched a rival VHS camcorder format, compatible with its VHS home video recorders but incompatible with Sony's 8 mm camcorders. JVC's VHS-C format received a big boost when several firms, including Matsushita, Toshiba, Hitachi, Mitsubishi, Philips, and Sharp, licensed the VHS-C technology from JVC.

In the period between 1985 and 1990, both the 8 mm and the VHS-C formats enjoyed fluctuating popularity in the market, and neither format gained market dominance. In addition, several other camcorder formats, including super VHS, super VHS-C, and digital 8 mm, were introduced in the early 1990s, each of which acquired niche market positions. Given this sequence of events, a dominant design did not emerge in the camcorder market at that time or since then, despite several firms having agreed on a standard two years before the commercial introduction of camcorders.

DVD Players (Dominant Design Emerged)

In an attempt to avoid the VHS–Betamax "format war" of the 1970s, in the mid-1990s, consumer electronics manufacturers, led by Sony, Toshiba, and Panasonic, in conjunction with movie studios, led by Warner and Columbia (a division of Sony), worked together in the DVD consortium. The objective of the DVD consortium was to establish the DVD standard as an open format. In addition, the DVDs were encoded with the Dolby Digital sound process to be compatible with virtually all home theater electronics.

However, the DVD format for DVD players was not unchallenged. In September 1997, Circuit City introduced a competing format called the Digital Video Express (DIVX), which was partially compatible with DVD. DIVX players could play all DVDs, but DVD players could not play DIVX discs, which were priced on par with DVD discs. Early adopters did not know it, but at the time of the DIVX announcement, Circuit City had neither the hardware nor the software for product demonstration.

When Circuit City finally launched DIVX in the fall of 1998, it faced an uphill battle, and studio support for DIVX never materialized. By May 1999, approximately two million DVD players had been shipped to retailers, compared with approximately 165,000 DIVX players. At the same time, there were 3317 DVD movie titles available on the DVD format, whereas only 471 titles were available on the DIVX format. In June 1999, Circuit City withdrew DIVX from the market, and the DVD format emerged as the dominant design.

REFERENCES

Abernathy, William J. and James M. Utterback (1978), "Patterns of Industrial Innovation," *Technology Review*, 8 (June–July), 40–47.

Adam, Jaffe, Manuel Trajtenberg, and Michael Fogarty (2000), "Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors," *American Economic Review*, 90 (2), 215–18.

Agarwal, Rajshree and Barry L. Bayus (2002), "The Market Evolution and Sales Take-Off of Product Innovations," *Management Science*, 48 (August), 1024–1041.

Ames, Edward and Nathan Rosenberg (1977), "Technological Change in the Machine Tool Industry, 1840-1910," in *Perspectives on Technology*, Nathan Rosenberg, ed. Cambridge, UK: Cambridge University Press

Amit, Raphael and Christopher Zott (2001), "Value Creation in E-Business," *Strategic Management Journal*, 22 (6–7), 493–520.

Anderson, Philip and Michael L. Tushman (1990), "Technological Discontinuities and Dominant Designs: A Cyclical Model of Technological Change," *Administrative Science Quarterly*, 35 (4), 604–633.

Augustin, Thomas (2005), "Correcting for Measurement Error in Parametric Duration Models by Quasi-Likelihood," working paper, Statistics Department, University of Munchen, Germany.

Basalla, George (1988), *The Evolution of Technology*. Cambridge, UK: Cambridge University Press.

Baum, Joel A.C., Helaine J. Korn, and Suresh Kotha (1995), "Dominant Designs and Population Dynamics in Telecommunications Services: Founding and Failure of Facsimile Transmission Service Organizations, 1965-1992," *Social Science Research*, 24 (2), 97–135.

Belsley D., E. Kuh, and R. Welsch (1980), *Regression Diagnostics*. Hoboken, NJ: John Wiley & Sons.

Bergen, Mark, Shantanu Dutta, and Orville C. Walker Jr. (1992), "Agency Relationships in Marketing: A Review of the Implications and Applications of Agency and Related Theories," *Journal of Marketing*, 56 (July), 1–24.

Besen, Stanley M. and Joseph Farrell (1994), "Choosing How to Compete: Strategies and Tactics in Standardization," *Journal of Economic Perspectives*, 8 (2), 117–31.

Brandenberger, Adam M. and Barry J. Nalebuff (1996), *Co-opetition*. Boston: Harvard Business School Press.

Chandy, Rajesh K. and Gerard J. Tellis (2000), "The Incumbent's Curse? Incumbency, Size, and Radical Product Innovation," *Journal of Marketing*, 64 (July), 1–17.

Christensen, Clayton M. (1997), *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Boston: Harvard Business School Press.

———, Fernando F. Suaréz, and James M. Utterback (1998), "Strategies for Survival in Fast-Changing Industries," *Management Science*, 44 (December), 207–220.

Cusumano, Michael, Y. Mylonadis, and Richard Rosenbloom (1992), "Strategic Maneuvering and Mass-Market Dynamics: The Triumph of VHS over Beta," *Business History Review*, 66 (1), 51–94.

David, Paul and Shane Greenstein (1990), "The Economics of Compatibility Standards: An Introduction to Recent Research," *Economics of Innovation and New Technology*, 1 (1–2), 3–41.

Dosi, Giovanni (1982), "Technological Paradigms and Technological Trajectories," *Research Policy*, 11 (3), 147–62.

Dyer, Jeffrey H. and Harbir Singh (1998), "The Relational View: Cooperative Strategy and Sources of Interorganizational Competitive Advantage," *Academy of Management Review*, 23 (4), 660–79.

Farrell, Joseph and Garth Saloner (1986), "Standardization and Variety," *Economics Letters*, 20 (March), 71–74.

——— and ——— (1988), "Coordination Through Committees and Markets," *Rand Journal of Economics*, 19 (2), 235–52

——— and ——— (1992), "Convertors, Compatibility, and the Control of Interfaces," *The Journal of Industrial Economics*, 60 (1), 9–35.

Frenken, Koen, Paolo P. Saviotti, and Michel Trommetter (1999), "Variety and Niche Creation in Aircraft, Helicopters, Motorcycles, and Microcomputers," *Research Policy*, 28 (5), 469–88.

Gallini, Nancy T. (1985), "Deterrence by Market Sharing: A Strategic Incentive for Licensing," *American Economic Review*, 74 (5), 931–41.

Garud, Raghu and Arun Kumaraswamy (1993), "Changing Competitive Dynamics in Network Industries: An Exploration of Sun Microsystems Open Systems Strategy," *Strategic Management Journal*, 14 (5), 351–69.

Goldenberg, Jacob, Barak Libai, and Eitan Muller (2002), "Is the Bandwagon Rolling? The Chilling Effect of Network Externalities on New Product Growth," working paper, Jerusalem School of Business Administration, Hebrew University of Jerusalem.

Golder, Peter N. (2000), "Historical Method in Marketing Research with New Evidence on Long-Term Market Share Stability," *Journal of Marketing Research*, 37 (May), 156–72.

——— and Gerard J. Tellis (1997), "Will It Ever Fly? Modeling the Takeoff of Really New Consumer Durables," *Marketing Science*, 16 (3), 256–70.

Greene, William (2002), *LIMDEP 8.0*. New York: Econometric Software Inc.

Grindley, Paul (1995), *Standards, Strategy and Policy*. New York: Oxford University Press.

Gupta, Sachin, Dipak C. Jain, and Mohanbir S. Sawhney (1999), "Modeling the Evolution of Markets with Indirect Network Externalities: An Application to Digital Television," *Marketing Science*, 18 (3), 396–416.

Iansiti, Marco and Jonathan West (1997), "Technology Integration: Turning Great Research into Great Products," *Harvard Business Review*, 75 (3), 69–79.

John, George, Allen M. Weiss, and Shantanu Dutta (1999), "Marketing in Technology-Intensive Markets: Toward a Conceptual Framework," *Journal of Marketing*, 63 (Special Issue), 78–91.

Katz, M. and Carl Shapiro (1986), "Technology Adoption in the Presence of Network Externalities," *Journal of Political Economy*, 94 (4), 822–41.

Khazam, Jonathan and David Mowery (1994), "The Commercialization of RISC: Strategies for the Creation of Dominant Designs," *Research Policy*, 23 (1), 89–102.

Krechmer, Ken (2000), "Fundamental Nature of Standards: Technical Perspective," *IEEE Communications Magazine*, 38 (6), 70–80.

Lawless, Jerald F. (2003), *Statistical Models and Methods for Lifetime Data*, 2d ed. Hoboken, NJ: John Wiley & Sons.

Lecocq, Xavier and Benoit Demil (2002), "Open Standard: Role of Externalities and Impact on Industry Structure," working paper, MIT Open Source Project.

Lee, J.P., D.E. Neal, M.W. Pruett, and H. Thomas (1995), "Planning for Dominance: A Strategic Perspective on the Emergence of a Dominant Design," *R&D Management*, 25 (1), 3–15.

Levin, Richard C., Alvin K. Klevorick, Richard R. Nelson, and Sidney. G. Winter (1987), "Appropriating the Returns from Industrial Research and Development," *Brookings Papers on Economic Activity*, No. 3, 783–820.

Lotz, Peter (1998), "The Paradox of High R&D and Industrial Stability: Technology and Structural Dynamics in the Global Hearing Instrument Industry," *Industry and Innovation*, 5 (2), 113–37.

Morrison, Donald G. (1969), "On the Interpretation of Discriminant Analysis," *Journal of Marketing Research*, 6 (May), 156–63.

National Science Foundation, Division of Science Resources Studies (1999), "Research and Development in Industry: 1997," NSF 99-358, (accessed June 10, 2004), [available at http://www.nsf.gov/sbe/srs/nsf99358/tables/nsf97a21.xls].

Rosenberg, Nathan (1994), *Exploring the Black Box: Technology, Economics and History*. Cambridge, UK: Cambridge University Press.

Schilling, Melissa (1998), "Technological Lockout: An Integrative Model of the Economic and Strategic Factors Driving Technology Success and Failure," *Academy of Management Review*, 23 (2), 267–84.

——— (2002), "Technology Success and Failure in Winner-Take-All Markets: The Impact of Learning Orientation, Timing, and Network Externalities," *Academy of Management Journal*, 45 (2), 387–98.

Schmidt, Peter and Ann D. Witte (1989), "Predicting Criminal Recidivism Using 'Split Population' Survival Time Models," *Journal of Econometrics*, 40 (1), 141–59.

Shankar, Venkatesh and Barry L. Bayus (2003), "Network Effects and Competition: An Analysis of the Home Video Game Industry," *Strategic Management Journal*, 24 (March), 375–84.

Shapiro Carl and Hal R. Varian (1999), "Art of Standard Wars," *California Management Review*, 41 (2), 8–32.

Sinha, Rajiv K. and Murali Chandrashekaran (1992), "A Split Hazard Model for Analyzing the Diffusion of Innovations," *Journal of Marketing Research*, 29 (February), 116–27.

Sirbu, Marvin and I. Zwimpfer (1985), "Computer Communication Standards: The Case of X. 25," *IEEE Communications*, 23 (3), 35–45.

Smith, Clayton G. (1997), "Design Competition in Young Industries: An Integrative Perspective," *Journal of High Technology Management Research*, 7 (2), 227–43.

Sorenson, Olav (2000), "Letting the Market Work for You: An Evolutionary Perspective on Product Strategy," *Strategic Management Journal*, 21 (5), 577–92.

Srinivasan, Raji, Gary L. Lilien, and Arvind Rangaswamy (2004), "First in, First out? Effect of Network Externalities on Pioneer Survival," *Journal of Marketing*, 68 (January), 41–57.

Stabell, Charles B. and Oystein D. Fjledstad (1998), "Configuring Value for Competitive Advantage: On Chains, Shops and Networks," *Strategic Management Journal*, 19 (5), 413–37

Suaréz, Fernando F. (2004), "Battles for Technological Dominance: An Integrated Framework," *Research Policy*, 33 (2), 271–86.

——— and James M. Utterback (1995), "Dominant Designs and the Survival of Firms," *Strategic Management Journal*, 16 (6), 415–30.

Teece, David J. (1986), "Profiting from Technological Innovations: Implications for Integration, Collaboration, Licensing and Public Policy," *Research Policy*, 15 (6), 285–305.

Tegarden, Linda F., Donald E. Hatfield, and Ann E. Echols (1999), "Doomed from the Start: What Is the Value of Selecting a Future Dominant Design?" *Strategic Management Journal*, 20 (6), 495–518.

Tsai, Diana H.A. (2005), "Knowledge Spillovers and High-Technology Clustering: Evidence from Taiwan's Hsinchu Science-Based Industrial Park," *Contemporary Economic Policy*, 23 (1), 116–28.

Tushman, Michael L. and Philip Anderson (1986), "Technological Discontinuities and Organizational Environments," *Administrative Science Quarterly*, 31 (3), 439–65.

——— and Johann Peter Murmann (1998), "Dominant Designs, Technology Cycles, and Organizational Outcomes," *Research in Organizational Behavior*, 20, 231–66.

——— and Lori Rosenkopf (1992), "Organizational Determinants of Technological Change: Toward a Sociology of Technological Evolution," *Research in Organizational Behavior*, 14, 311–47.

Utterback, James M. (1994), *Mastering the Dynamics of Innovation: How Companies Can Seize Opportunities in the Face of Technological Change*. Boston: Harvard Business School Press.

Van Bruggen, Gerritt H., Gary L. Lilien, and Manish Kacker (2002), "Informants in Organizational Marketing Research: Why Use Multiple Informants and How to Aggregate Responses," *Journal of Marketing Research*, 39 (November), 469–78.

Van de Ven, Andrew and Raghu Garud (1993), "Innovation and Industry Development: The Case of Cochlear Implants," in *Research on Technological Innovation, Management and Policy*, Vol. 5, R. Burgelman and R. Rosenbloom, eds. Greenwhich, CT: JAI Press, 1–46.

Wade, James (1995), "Dynamics of Organizational Communities and Technological Bandwagons: An Empirical Investigation of Community Evolution in the Microprocessor Market," *Strategic Management Journal*, 16 (5), 111–33.

The Wall Street Journal (2004), "IBM, Intel to Unveil Blade Plan," (September 2), (accessed September 6, 2004), [available at http://www.wsj.com].

West, Joel (2003), "How Open Is Open Enough? Melding Proprietary and Open Source Platform Strategies," *Research Policy*, 32 (7), 1259–85.

Wilson, Lynn O., Allen M. Weiss, and George John (1990), "Unbundling of Industrial Systems," *Journal of Marketing Research*, 27 (May), 123–38.

Franziska Völckner & Henrik Sattler

Drivers of Brand Extension Success

The research presented in this article addresses the issue of the significance and relative importance of the determinants of extension success by simultaneously investigating ten success factors. The empirical analysis considers the direct relationships between success factors and extension success, the structural relationships among investigated factors, and moderating effects. The authors find that fit between the parent brand and an extension product is the most important driver of brand extension success, followed by marketing support, parent-brand conviction, retailer acceptance, and parent-brand experience. The authors also find several important structural relationships among the investigated success factors (e.g., marketing support → fit → retailer acceptance → extension success). Finally, the interaction terms of fit with the quality of the parent brand and with parent-brand conviction are statistically significant, albeit of relatively low importance.

Brand extensions—that is, use of established brand names to launch new products (e.g., use of the name Mr. Clean for a new detergent)—represent one of the most frequently used branding strategies. Extending brands both within and beyond the original product category is deemed to be profitable because, in general, it is assumed that brands that are already known and recognized require lower new product introduction expenses, such as advertising, trade deals, or price promotions (Collins-Dodd and Louviere 1999; Tauber 1988). Nevertheless, the success of brand extensions is uncertain. Failure rates of brand extensions in many fast-moving consumer good (FMCG) product categories are approximately 80% (Ernst & Young and ACNielsen 1999; *Marketing* 2003). Therefore, potential determinants of brand extension success (i.e., success factors of brand extensions) have emerged as an important focus of research inquiry to provide insights that may help managers reduce the failure rates of brand extensions (e.g., Aaker and Keller 1990; Bottomley and Doyle 1996; Dacin and Smith 1994; Swaminathan, Fox, and Reddy 2001). The findings of previous studies provide important insights into the factors that influence brand extension success. In particular, they reveal the kinds of success factors that play an important role in the success of an extension product, at least under certain conditions. For example, despite some mixed results, prior studies show that the quality of the parent brand and the fit between the parent brand and extension product categories are highly important brand extension success factors (Bottomley and Holden 2001). Overall, approximately 15 determinants of extension success have been proved significantly relevant ($p < .10$) in at least one empirical study (for a more detailed discussion, see the subsection "Determinants of Brand Extension Success" in the "Conceptual Framework" section).

However, there are at least two key issues that have received little or no attention in prior work: First, little is known about the relative importance of the success factors in explaining brand extension success because each previous study investigated the effects of only a small fraction of all relevant success factors at one time (usually two to four factors). Second, previous studies tested only the direct relationship between brand extension success (dependent variable) and potential success factors (independent variables). They did not take into account that some success factors may constitute dependent variables in other structural relationships; that is, previous studies did not examine a series of structural relationships. The failure to account for potential relationships among success factors may cause faulty interpretation of the significance and relative importance of the success factors under investigation.

Against this background, this article makes two primary contributions to brand extension research by presenting a large-scale empirical study: First, we address the issue of the significance and relative importance of the determinants of extension success by simultaneously investigating ten success factors. Second, we apply a structural equation analysis to test several conceptual models of the determinants of extension success. Our analysis considers the direct relationships between success factors and extension success, the structural relationships among investigated factors, and moderating effects. Furthermore, an advantage of structural equation analysis is the ability to account for measurement errors. The contributions of our study are especially relevant for marketing practice. To improve brand extension success, it is imperative for managers to know which of the large number of potentially relevant success factors should receive the most attention and how they should allocate resources to the relevant factors.

We consider brand extensions within a comprehensive nomological net by developing a conceptual framework for our empirical research that builds on the results of prior research and the managerial wisdom derived from expert judgments. We specify and empirically test four alternative models. The focus of this article is not on theory development; rather, we pursue an empirical approach to studying the success determinants of brand extensions by unifying

---

Franziska Völckner is Assistant Professor of Marketing and Management (e-mail: voelckne@econ.uni-hamburg.de), and Henrik Sattler is Professor of Marketing and Management (e-mail: uni-hamburg@henriksattler.de), Institute of Marketing and Retailing, University of Hamburg. The authors are grateful to John Roberts and the four anonymous *JM* reviewers for helpful comments on previous versions of this article.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

findings from published research and beliefs of managers into a comprehensive model of brand extension success.

Having outlined the motivation behind this article, in the next section, we develop the conceptual framework for our empirical research. Then, we describe the research design and discuss some methodological issues. We present the results of our study and conclude our article by outlining the implications of the findings, limitations of the study, and opportunities for further research.

Conceptual Framework

Overview of the Conceptual Framework

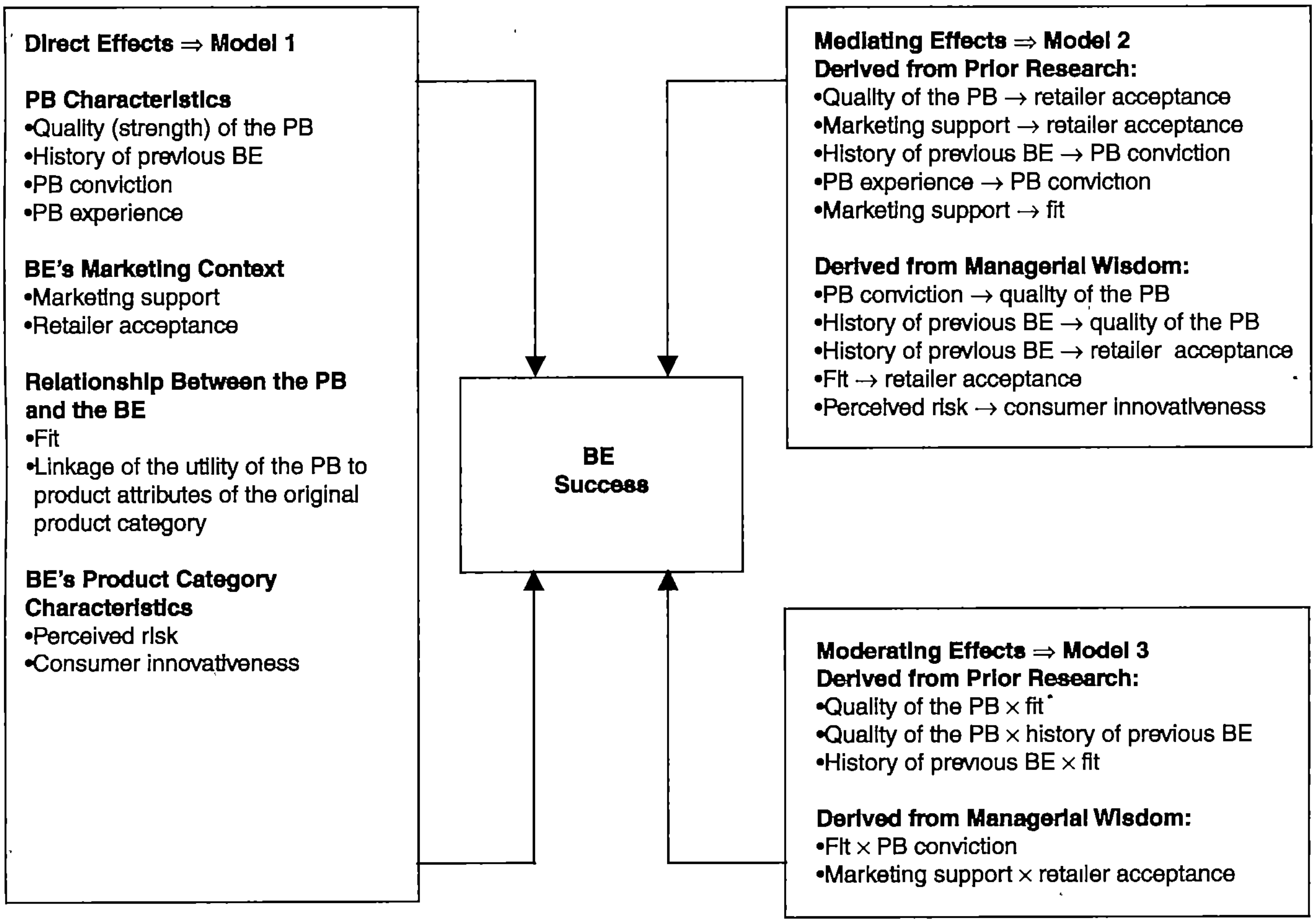
Our conceptual framework identifies potentially relevant success factors and specifies the system of relationships among these factors. Although we refer to the various relationships as hypotheses, this article focuses on testing the fit of alternative models and does not focus on testing various specific hypotheses. We opted to use the term "hypotheses" for the relationships in the models because it is the common term used in the literature.

Figure 1 provides an overview of our conceptual framework. The potentially relevant determinants of brand extension success can be assigned to four groups: (1) parent-brand characteristics, (2) the extension's marketing context, (3) the relationship between the parent brand and the extension product, and (4) the extension's product category characteristics. Figure 1 proposes that the success of a brand extension is influenced by the direct effects of determinants, mediating effects, and moderating effects. The following discussion presents the details of the constituent elements of our framework.

Determinants of Brand Extension Success

This section presents the results of a systematic literature review that resulted in identification of a broad variety of determinants of extension success.[1] The review included the following journals between 1985 and 2001 (the year of data

FIGURE 1
Overview of Conceptual Framework


Direct Effects ⇒ Model 1
PB Characteristics
•Quality (strength) of the PB
•History of previous BE
•PB conviction
•PB experience
BE's Marketing Context
•Marketing support
•Retailer acceptance
Relationship Between the PB and the BE
•Fit
•Linkage of the utility of the PB to product attributes of the original product category
BE's Product Category Characteristics
•Perceived risk
•Consumer innovativeness
BE Success
Mediating Effects ⇒ Model 2
Derived from Prior Research:
•Quality of the PB → retailer acceptance
•Marketing support → retailer acceptance
•History of previous BE → PB conviction
•PB experience → PB conviction
•Marketing support → fit
Derived from Managerial Wisdom:
•PB conviction → quality of the PB
•History of previous BE → quality of the PB
•History of previous BE → retailer acceptance
•Fit → retailer acceptance
•Perceived risk → consumer innovativeness
Moderating Effects ⇒ Model 3
Derived from Prior Research:
•Quality of the PB × fit
•Quality of the PB × history of previous BE
•History of previous BE × fit
Derived from Managerial Wisdom:
•Fit × PB conviction
•Marketing support × retailer acceptance

Notes: PB = parent brand; BE = brand extension. Model 1: direct effects; Model 2: direct and mediating effects; Model 3: direct, mediating, and moderating effects.

---

[1] Our research does not consider the effect of ingredient branding strategies that has been investigated in recent empirical studies (e.g., Desai and Keller 2002).

collection): *International Journal of Research in Marketing, Journal of Consumer Research, Journal of Marketing, Journal of Marketing Research, Management Science, Marketing Letters*, and *Marketing Science*. In addition, we included articles that were frequently cited in these journals in the review. We limited our review to empirical studies that investigate the effect of certain success factors (e.g., quality of the parent brand) on some kind of extension success measure (typically, customer evaluations of brand extensions linked, for example, to the perceived quality of the extension). Most of the studies were based on consumer surveys that used hypothetical brand extensions (i.e., extensions not introduced to the market). Respondents were asked to rate independent (i.e., success factors) and dependent (i.e., success of the extension; e.g., the perceived quality of the extension) variables on rating scales. Aaker and Keller's (1990) fundamental study represents a typical example. Few studies used market data as dependent variables (e.g., market share or stock market value; Lane and Jacobson 1995; Reddy, Holak, and Bhat 1994; Smith and Park 1992).

We identified 15 determinants of extension success that had been proved significantly ($p < .10$) relevant in at least one empirical study. We excluded the following three determinants because they had only minor influence on brand extension success in prior studies: (1) difficulty in producing a product from the extension's product class (e.g., Bottomley and Doyle 1996; Bottomley and Holden 2001; Nijssen and Bucklin 1998; Sunde and Brodie 1993), (2) consumers' knowledge of the extension's product class (e.g., Smith and Park 1992), and (3) company size (Reddy, Holak, and Bhat 1994). Likewise, we excluded the order and direction of previous brand extensions (Dawar and Anderson 1994) and consumers' mood (Barone, Miniard, and Romeo 2000) because measuring these two factors is extremely complex within the scope of a consumer survey. Furthermore, on the basis of the results of a comprehensive pretest, we found that only a few consumers could correctly state the order and direction of previous brand extensions of real brands. This finding suggests that the order and direction of previous extensions play only a minor role in consumers' evaluations of a new brand extension product.

In summary, the literature review resulted in ten success factors that may influence brand extension success. These factors are included in our empirical study (see the second column of Table 1). Given that we considered the results of 45 empirical studies that cover 15 years of extensive research effort, it seems reasonable to assume that the ten identified success factors cover a substantial subset of all potentially relevant success factors. The term "potentially relevant" seems adequate because there are many examples of mixed results for a particular factor (e.g., Dacin and Smith [1994] and Reddy, Holak, and Bhat [1994] find mixed results on the number of previous brand extensions). In other words, a certain factor that was significant in one study was not significant in another study.

In addition to prior work, interviews with brand managers and researchers can also be useful in broadening the scope of knowledge about brand extension evaluations. Therefore, we used expert knowledge as an additional information source to build further confidence in the relevance of identified brand extension success factors. There is an important tradition in marketing research of incorporating judgmentally derived information into marketing models (Leeflang and Wittink 2000), especially managerially estimated parameter values (e.g., Sandor and Wedel 2001). We interviewed 21 brand managers and researchers. We set the requirement that all experts needed to be highly knowledgeable in brand extension issues and in the German FMCG market. All brand managers hold senior management positions in major German FMCG companies. The researchers stem from marketing departments at several German universities. We conducted 21 semistructured interviews regarding the relevance of success factors and the system of relationships among them. During the interviews, we gave participants a description of the ten success factors, which we identified in the literature review, and asked them to evaluate the factors' relevance. Almost all experts agreed that, in general, the ten success factors are relevant for brand extension success. Finally, we asked the experts to add success factors that might play a major role for brand extension success. Most of them did not name any additional factors. Within the group of experts that mentioned additional determinants of extension success, only one type of factor was mentioned more than once, namely, profitability of the market as one of the characteristics of the extension's market. Although it is highly plausible to assume that this factor is important for new product success, we decided not to explore it in our study, because our analysis focused on consumers' evaluations of brand extensions. We also needed to restrict the number of success factors in our study to keep the evaluation task manageable for respondents.

Table 1 (H_1–H_{10}) presents the success factors that we considered in our study and their corresponding hypotheses. We abridged the original wording of the hypotheses because of limited space. We treat multiple measures of the same factor (e.g., fit or history of previous extensions) as one determinant of extension success. Table 1 quotes only one source per hypothesis. The sources listed provide a theoretical foundation for the postulated hypotheses. Therefore, we dismissed a detailed discussion of the ten direct effects. However, we provide a detailed explanation of the mediating and moderating effects because not all of them build on the results of prior research.

Mediating Effects

Mediating effects derived from prior research. Prior research has focused on the relationship between extension success and its potential determinants. On the basis of our systematic literature review, we identified four studies (see Table 1) that investigated the impact of a certain success factor (independent variable) on another success factor (dependent variable). Nevertheless, the issue of how structural relationships among the success factors affect extension success is still not clear, because none of the studies considered extension success a component of the estimated model.

To begin with, retailers' utility for brand names is based on associations, such as retail performance expectations and positioning characteristics. Strong brands have more posi-

TABLE 1
Hypotheses Regarding Success Factors of Brand Extensions

	A Brand Extension Is More Successful ...	Source
Main Effects: Direct Effects of the Success Factors		
Parent-Brand Characteristics		
H_1 Quality (strength) of the parent brand	If the quality of the parent brand is high.	Smith and Park (1992)
H_2 History of previous brand extensions	If the history of previous brand extensions is successful. For example,	
	•High number of previous brand extensions.	Dacin and Smith (1994)
	•High variability among product types offered by the parent brand.	Boush and Loken (1991)
	•Low variance in quality among previous brand extensions.	Dacin and Smith (1994)
H_3 Parent-brand conviction	If parent-brand conviction is high.	Kirmani, Sood, and Bridges (1999)
H_4 Parent-brand experience	If parent-brand experience is high.	Swaminathan, Fox, and Reddy (2001)
Extension's Marketing Context		
H_5 Marketing support	If the marketing support is high. For example,	
	•Advertising support.	Reddy, Holak, and Bhat (1994)
	•Firm's marketing competence.	Reddy, Holak, and Bhat (1994)
H_6 Retailer acceptance	If the retailer acceptance is high.	Nijssen (1999)
Relationship of Parent Brand to Extension Product		
H_7 Fit between parent brand and extension product	If the fit between the parent brand and the extension is high. For example,	
	•High global similarity.	Aaker and Keller (1990)
	•High ability of the owner of the parent brand to make a product in the extension product class.	Aaker and Keller (1990)
	•High relevance of the extended associations for the extension product.	Broniarczyk and Alba (1994)
H_8 Linkage of the utility of the parent brand to product attributes of the original product category	The less the utility of the parent brand is linked to product attributes of the original product category.	Rangaswamy, Burke, and Oliva (1993)
Extension's Product Category Characteristics		
H_9 Perceived risk	If the perceived risk is low.	Nijssen and Bucklin (1998)
H_{10} Consumer innovativeness	If consumer innovativeness is high.	Klink and Smith (2001)

	Expected Sign	Source
Mediating Effects: Dependence Relationships Between Success Factors		
H_{11} Quality of the parent brand → retailer acceptance	+	Collins-Dodd and Louviere (1999)
H_{12} Marketing support → retailer acceptance	+	Collins-Dodd and Louviere (1999)
H_{13} History of previous brand extensions → parent-brand conviction	+	DelVecchio (2000)
H_{14} Parent-brand experience → parent-brand conviction	+	Kirmani, Sood, and Bridges (1999)
H_{15} Marketing support → fit between parent brand and extension product	+	Klink and Smith (2001)
Moderating Effects		
H_{21} Quality of the parent brand × fit	+	Aaker and Keller (1990)
H_{22} Quality of the parent brand × history of previous brand extensions	+	Keller and Aaker (1992b)
H_{23} History of previous brand extensions × fit	+	Dacin and Smith (1994)

tive associations of sales volume, competitive clout, and traffic-generating appeal. Retailers also expect that manufacturers are unlikely to place a strong brand name on an inferior product and risk diluting its brand equity. Therefore, the quality or strength of a parent brand is expected to have a positive influence on retailer acceptance of a new extension product (H_{11}; Collins-Dodd and Louviere 1999). Furthermore, in general, retailers appreciate manufacturers' consumer advertising because it increases consumers' product awareness, thereby preselling the product and reducing retailers' selling costs. Therefore, greater marketing support in terms of consumer advertising should generate more retailer acceptance (H_{12}; Collins-Dodd and Louviere 1999). Prior research has also found support for a positive impact of the history of previous brand extensions and parent-brand experience on parent-brand conviction. It might be expected that consumers' exposure to the brand name increases as the number of previous brand extensions increases. If we assume that experiences are positive, greater exposure to the brand name may generate greater liking for and trust in the brand name.[2] Therefore, the history of previous extensions might have a positive effect on brand conviction (H_{13}; DelVeechio 2000). Related, direct experience with the parent brand may increase the personal relevance of the brand and generate greater liking for the brand. This phenomenon has been labeled the "mere ownership effect." Direct experience might also cause greater parent-brand knowledge, stronger brand associations, and stronger autobiographical memories, all of which might result in higher levels of parent-brand conviction (H_{14}; e.g., Kirmani, Sood, and Bridges 1999). Finally, research on typicality suggests that repeated exposure to an extension product may elevate consumers' perceptions of fit. Greater exposure to the brand extension helps consumers identify more shared attributes between the parent brand and the extension. Thus, higher levels of marketing support in terms of advertising lead to higher levels of perceived fit (H_{15}; e.g., Klink and Smith 2001; Lane 2000).

Mediating effects derived from managerial wisdom. During the interviews with brand managers and researchers, we gave them a description of the ten success factors and asked them to point out possible structural relationships among the factors and to give a reason for each. The experts specified a total of 13 structural relationships. However, just one or a few respondents pointed out several of these relationships. We decided to exclude the relationships that were specified by only five or less experts (25%) to avoid any influence of personal subjective opinions on model development. As a result, we identified 8 structural relationships that were specified by at least six or more participants (see Table 2, Panel A).

In general, the experts provided highly plausible reasons for the hypothesized relationships among success factors. A total of 13 participants expected a positive effect of parent-brand conviction on perceived quality of the parent brand (H_{16}). They reasoned that consumers who like or trust a parent brand should have more favorable beliefs about brand attributes and greater confidence in their beliefs than con-

TABLE 2
Judgmentally Derived Mediating and Moderating Effects of the Success Factors

A: Personal Interviews with Brand Managers and Researchers

Mediating Effects: Dependence Relationships Between Success Factors	Expected Sign	Percentage of Experts
H_{16} Parent-brand conviction → quality of the parent brand	+	62
H_{17} History of previous brand extensions → quality of the parent brand	+	52
H_{18} History of previous brand extensions → retailer acceptance	+	38
H_{19} Fit between parent brand and extension → retailer acceptance	+	29
H_{20} Perceived risk of unknown brands → consumer innovativeness	–	29
See Table 1		
H_{11} Quality of the parent brand → retailer acceptance	+	48
H_{12} Marketing support → retailer acceptance	+	43
H_{15} Marketing support → fit between parent brand and extension product	+	38

B: Group Discussion with Brand Managers and Researchers

Moderating Effects	Expected Sign
H_{24} Fit × parent-brand conviction	+
H_{25} Marketing support × retailer acceptance	+
See Table 1	
H_{21} Quality of the parent brand × fit	+
H_{22} Quality of the parent brand × history of previous brand extensions	+
H_{23} History of previous brand extensions × fit	+

---

[2] Although negative brand experiences that make consumers dislike the brand are possible, such experiences are likely to be rare; otherwise, brands would not survive in the marketplace.

sumers who dislike the brand. Moreover, strong and favorable brand associations enable the brand to be strategically differentiated and positioned in the consumer's mind. High levels of parent-brand conviction reflect favorable predispositions toward the brand and may function as a "filter" for how a consumer evaluates the specific attributes of the brand (including perceived quality), all of which would lead to a positive relationship between parent-brand conviction and parent-brand quality. Within this context, 11 experts also expected a positive influence of the history of previous brand extensions on perceived quality of the parent brand (H_{17}). It might be expected that consumers' exposure to the brand name and their likelihood to try one of the brand's products increase as the number of previous brand extensions increases. Product trial provides new information about a brand. Information generated from product trial typically results in increased brand recall and stronger brand attitudes, which in turn have a powerful and positive impact on parent-brand evaluation, unless product experience is negative (Swaminathan, Fox, and Reddy 2001). Moreover, several experts agreed that the history of previous brand extensions has a positive impact on retailer acceptance of a new extension product (H_{18}). They assumed that a successful history of previous extensions typically results in favorable beliefs about extension attributes, such as retail performance expectations and positioning characteristics. They also assumed that a successful history of brand extensions might increase retailers' confidence in and favor of their evaluations of subsequent brand extensions. Retailers may perceive the parent brand as highly competent in making and introducing extension products, which would also result in a positive relationship between the history of previous extensions and retailer acceptance. Continuing with possible effects on retailer acceptance, 6 participants expected a positive impact of perceived fit on retailer acceptance (H_{19}). Retailers want to avoid creating poor assortment perceptions among customers. Not listing a new extension product, which is obviously affiliated with the parent brand (in terms of perceived fit between the extension product and the parent brand), may create the perception of an incomplete assortment. Therefore, high levels of fit might enhance the perceived importance of listing the new extension product to maintain store selection and quality reputations, which initiates a positive relationship between fit and retailer acceptance. Finally, 6 experts expected that the perceived risk of unknown brands would have a negative impact on consumer innovativeness (H_{20}). They believed that the desire or willingness to try new experiences should decrease as the level of perceived risk increases. Diffusion theory holds that people respond differently to new products. To some extent, the response differences reflect differences in risk-taking propensity. Later adopters tend to be more risk averse than earlier adopters (Rogers 1983), which supports the notion that high levels of perceived risk are associated with a low desire to try new products.

Moderating Effects

Moderating effects derived from prior research. In addition to the ten main effects variables (H_1–H_{10}, Table 1) and the ten indirect effects (H_{11}–H_{20}, Tables 1 and 2), a third group of effects, namely, moderating or interaction terms, may play an important role in determining extension success. Given the results of prior studies, there are three interactions between the success factors under investigation that are expected to influence extension success (see Table 1, H_{21}–H_{23}).

The premise for extending an existing brand name is that consumers use their beliefs about the parent brand to draw conclusions about an extension product. The degree to which brand associations are transferred to the extension depends on the level of perceived fit between the brand and the extension product. Therefore, a positive interaction is expected between the fit variable and the quality of the parent brand (e.g., Aaker and Keller 1990); that is, the positive effect of the quality of the parent brand on extension success increases as the level of perceived fit increases (H_{21}). Consumers also tend to be much more confident about their judgments of parent-brand quality if they are based on a large sample of instances (i.e., products that bear the parent brand name) rather than on a comparably small one. The "database" from which consumers draw information for forming judgments about the brand extension increases with the increase in the number of products affiliated with the brand (e.g., Dacin and Smith 1994). Consequently, the positive effect of higher levels of parent-brand quality on extension success increases as the number/success of previous brand extensions increases (H_{22}). Consumers' skepticism arising from low fit can also be thwarted if a company has historically demonstrated that it provides consistently good products across several product categories (e.g., Dacin and Smith 1994). Therefore, the negative effect of low levels of fit on extension evaluations diminishes as the number/success of previous brand extensions increases (H_{23}).

Moderating effects derived from managerial wisdom. We used expert knowledge to gain further insight into potentially relevant interaction terms. We acquired expert knowledge by conducting a second expert survey during a group discussion. The group consisted of three senior brand managers (from major German FMCG companies) and three researchers (from marketing departments of German universities). We presented the experts with ten success factors that we identified in the literature and asked them to identify highly relevant interaction effects between these factors. The six experts were required to come up with a solution during the group discussion that represented a consensus of all their opinions. This procedure resulted in five interaction terms: the aforementioned moderating effects (H_{21}–H_{23}), an interaction between fit and parent-brand conviction (H_{24}), and an interaction between marketing support and retailer acceptance (H_{25}; see Table 2, Panel B). The experts provided highly plausible reasons for the interaction terms. For example, the degree to which brand trust and likeability are transferred to an extension product depends on the level of perceived fit between the brand and the extension, similar to the interaction between fit and the quality of the parent brand. Higher levels of perceived fit allow for better transfer of brand associations (i.e., affect and beliefs). Therefore, a positive interaction between the fit variable and parent-brand conviction is expected (H_{24}). Likewise, the experts agreed that the positive effects of high levels of consumer advertising on extension success depend (to an extent) on the distribution support of the extension (H_{25}). Consumer advertising is expected to increase consumers' product awareness and utility, thus generating consumer pull. The degree to which increased awareness and product utility translate into purchases of the extension product depends (to an extent) on the availability of the item on the shelf. Thus, consumer advertising positively influences extension success, provided that the product is available in retail stores, which brings about a positive interaction between the extension's marketing support and retailer acceptance.

Model Specification

As we previously mentioned, the primary purpose of this study is to determine empirically the mediating and moderating effects between a broad set of determinants of brand extension success and the extent to which these determinants affect the success of brand extensions. Therefore, we integrated the identified direct, indirect, and moderating effects into alternative models. First, we estimated a main effects–only model (M1). Model M1 contains the direct effects of success factors on extension success that have been proposed in academic literature (H_1–H_{10}; see Table 1). Second, we specified Model M2, which considers the direct effects we investigated in M1 and the identified structural relationships among the determinants of extension success (H_1–H_{20}; see Tables 1 and 2). Third, we introduced moderator effects in Model M3, which contains the direct and indirect effects we investigated in M2 and the identified moderator effects (H_1–H_{25}). We included the moderator variables by using a step-by-step procedure that resulted in two significant moderator effects (M3a and M3b). Finally, Model M4 explores the moderating effect of consumers' familiarity with the extension product by means of a two-group analysis.

Research Design

We tested the proposed models using a broad variety of real extensions and real parent brands in the German FMCG industry. We conducted two pretests to identify the parent brands and extension products that we finally used in the main study. We began with a set of 50 brands, taken from a list of the 100 brands on the German FMCG market, that generate the highest sales revenues. We limited our sample of brands to those that had launched at least five extensions within the past four years. We asked respondents (n = 40) in the first pretest to rate their familiarity with these brands. We chose the 25 brands with the highest awareness ratings. We concentrated on well-known brands in our study because brand extensions are particularly popular in the marketplace for well-established brands (e.g., Swaminathan, Fox, and Reddy 2001; Tauber 1988).

We randomly selected five real brand extensions introduced between 1998 and 2001 for each of the 25 brands from the German LPV New Products Database.[3] (When we designed our study, the LPV database contained all the new products introduced in Germany since January 1996.) We conducted a second pretest with a convenience sample of 171 participants to warrant that the brand extensions finally used in our study varied in terms of awareness and (perceived) extension success. The participants answered a series of questions about their familiarity with the extension products and their perceived quality of the extensions. We selected 22 parent brands with three extensions each on the basis of the pretest results. This process yielded 66 brand extensions for the main study.

We used multiple-item measures for extension success and for success factors. We developed a questionnaire using variables and rating scales that closely coincide with those used in previous studies (see the Appendix). We conceptualized extension success as perceived extension quality (e.g., Aaker and Keller 1990; Bottomley and Doyle 1996; Broniarczyk and Alba 1994; Klink and Smith 2001) and measured it with three questions that ascertained (1) the perceived overall quality of the extension (–3 = "very low," and +3 = "very high"), (2) the market position of the extension product (extent to which participants agreed/disagreed with the statement "[extension product] belongs to the top 3 products in the market" [–3 = "strongly disagree," and +3 = "strongly agree"]), and (3) the perceived overall quality relative to competing products in the market (–3 = "below average," and +3 = "above average").

We took measures of success factors from typical previous studies. For example, we measured the perceived fit by asking participants to rate the following kinds of items: (1) the overall similarity of the brand extension to the parent brand (–3 = "not very similar," and +3 = "very similar"), (2) the perceived ability of the company to make a product in the extension product class ("Would the people, facilities, and skills used in making the original product be helpful if the manufacturer were to make the extension product?" [–3 = "not at all helpful," and +3 = "very helpful"]), and (3) the relevance of the brand-specific associations in the extension product category (Broniarczyk and Alba 1994). The Appendix provides an overview of all the measures we used.

The questionnaire was administered to people living in Germany by means of door-to-door interviews. Respondents participated voluntarily without any compensation and were selected according to a quota sampling procedure with a representative structure for Germany in terms of age, gender, and number of household members. In addition, we required respondents to know the brands that appeared on the questionnaire. Students majoring in marketing conducted the interviews. They attended detailed training sessions before going into the field. Interviewers obtained respondents' names and telephone numbers for validation purposes. We randomly contacted 5% of the respondents to confirm that the interviews were completed as planned. A total of 2426 respondents completed the questionnaire, and 7278 cases were obtained because each questionnaire had scales for 3 of the 66 brand extensions (2426 × 3 = 7278). We excluded respondents for whom extension categories were not relevant, which reduced the number of cases to 6668. We used a combination of listwise deletion and the informative imputation procedure (Ford 1976) to handle missing values (.2%).

---

[3] The initialism LPV stands for "Lebensmittelpraxis Verlag," an information service for the German FMCG industry.

We used STREAMS 2.50 (Gustafsson and Stahl 2000) to handle the hierarchical structure (i.e., "repeated measures" nature) of our data, applying two-level structural equation modeling within a framework of maximum likelihood estimation (e.g., Muthén 1989, 1990). STREAMS provides an interface to the estimation program AMOS 4.0 (Arbuckle and Wothke 1999).[4] We performed outlier analysis and excluded influential outliers before we estimated each of the specified structural equation models (e.g., Baumgartner and Homburg 1996, p. 148; Bollen 1989; Mullen, Milne, and Doney 1995).

The sufficiently large sample enabled us to split the data randomly into two parts (Sample 1 and Sample 2), facilitating the cross-validation of findings (Byrne 2001). Sample 1 served as the calibration sample, which we used to test the hypothesized models. We subsequently tested the replicability of each model on the basis of Sample 2 (the validation sample).

Results

We present the results of our study in three sections. First, we briefly describe the results of the assessment of the multiple-item measures we used. Second, we analyze different models with the direct, indirect, and moderating effects of the ten success factors shown in Tables 1 and 2. Third, we illustrate the managerial relevance of our findings by applying our "best" model to several hypothetical extension decisions and comparing it with an intuitive model advocated by managers.

Measurement Model Assessment

In our examination of the reliability and validity of our measures, we computed Cronbach's alpha for all scales. The alphas varied from .80 to .92. We then subjected the items to exploratory and confirmatory factor analyses. We ascertained the quality of the construct measurement by examining the sign, size, and significance of the estimated factor loadings and the magnitude of measurement error. All the factor loadings were significant (t > 30.54), and they were highly related to their respective constructs, in support of convergent validity. Explained variance ranged from 52.9% to 84.9%. We computed two reliability indexes to ascertain how well the constructs are measured by their indicators (e.g., Bagozzi and Yi 1988; Fornell and Larcker 1981; Steenkamp and Van Trijp 1991). All constructs displayed satisfactory levels of reliability, as indicated by individual-item reliabilities in excess of .70 and composite reliabilities ranging from .62 to .85. These figures are comparable to figures obtained in many marketing studies (e.g., Edison and Geissler 2003; Laroche, Bergeron, and Goutaland 2001). Finally, we conducted a test for construct discriminant validity. We calculated the shared variance estimate for each success factor and compared it with the square of the phi coefficient that represents the correlation between pairs of success factors (Fornell and Larcker 1981). Each of the shared variance estimates exceeded the square of the corresponding phi coefficient, which provided evidence of discriminant validity among all success factors. They represent distinct constructs in the structural equation models. In summary, the measures fit well with the data, and we used them in the following process to estimate and test the models previously described.

Estimating and Testing Models with Direct, Indirect, and Moderating Effects

Table 3 reports the parameter estimates and goodness-of-fit indicators of Models M1, M2, and M3. We use overall fit indicators to assess how well the models fit the data (e.g., Hu and Bentler 1995). All fit indicators are within a satisfactory range and show that the models account for a substantial amount of variance.

A comparison of the path coefficients across models denotes that there are almost no differences, indicating that the estimates are robust. Furthermore, we cross-validated our findings by comparing the estimates across the calibration and the validation sample as a means to assess the robustness of the postulated model. The test involved constraining the path coefficients and factor loadings and the variance–covariance of the constructs to be equal across the two samples. We then evaluated whether the resulting change in chi-square was significant ($p < .01$), with the appropriate number of degrees of freedom. We observed no statistically significant differences. This indicates that the models are robust when other samples are used.

Effects of parent-brand characteristics. All the models provided statistically significant support for the positive role of the parent brand's quality and consumers' brand conviction in determining brand extension success. We found notable differences for the history of previous brand extensions and parent-brand experience between the model that reflected only direct effects (M1) and the models that included indirect (M2) and moderator (M3) effects. When we examined only direct effects, parent-brand experience had no influence on extension success. Nevertheless, M2 and M3 reveal that there is a powerful and highly significant relationship between parent-brand experience and parent-brand conviction (standardized path estimate = .64, $p < .01$). In other words, parent-brand experience has a significant and positive impact on extension success due to its positive effect on parent-brand conviction, even in the absence of any significant direct relationship between parent-brand experience and extension success. Therefore, ignoring the indirect effects of a variable results in faulty interpretation of the relevance of this factor for brand extension success. In this case, the estimated influence of the parent-brand experience on the success of a brand extension increases from .00% (nonsignificant direct effect in M1) to 10.39% (M2).

Model M1 similarly provides a biased estimate for the effect of the history of previous brand extensions on brand extension success. In M1, the direct influence of the history

---

[4] Covariance matrices are available on request.

TABLE 3
Structural Parameter Estimates and Goodness-of-Fit Indicators

Paths		Estimates			
		M1	M2	M3a	M3b
H_1	Quality of the parent brand → extension success	+.10**	+.11**	+.11**	+.11**
H_2	History of previous brand extensions → extension success	–.04*	n.s.	n.s.	n.s.
H_3	Parent-brand conviction → extension success	+.25**	+.30**	+.30**	+.30**
H_4	Parent-brand experience → extension success	n.s.	n.s.	n.s.	n.s.
H_5	Marketing support → extension success	+.13**	+.14**	+.13**	+.13*
H_6	Retailer acceptance → extension success	+.29**	+.30**	+.30**	+.30**
H_7	Fit → extension success	+.45**	+.46**	+.47**	+.47**
H_8	Linkage of the utility of the parent brand to specific product attributes → extension success	n.s.	n.s.	n.s.	n.s.
H_9	Perceived risk of unknown brands → extension success	+.08**	+.09**	+.08**	+.08**
H_{10}	Consumer innovativeness → extension success	+.05**	+.06**	+.06**	+.07**
H_{11}	Quality of the parent brand → retailer acceptance	—	n.s.	n.s.	n.s.
H_{12}	Marketing support → retailer acceptance	—	+.56**	+.56**	+.56**
H_{13}	History of previous brand extensions → parent-brand conviction	—	.13**	.13**	.13**
H_{14}	Parent-brand experience → parent-brand conviction	—	.64**	.64**	.64**
H_{15}	Marketing support → fit	—	+.40**	+.40**	+.40**
H_{16}	Parent-brand conviction → quality of the parent brand	—	+.74**	+.74**	+.74**
H_{17}	History of previous brand extensions → quality of the parent brand	—	+.05*	+.04*	+.04*
H_{18}	History of previous brand extensions → retailer acceptance	—	+.11**	+.11**	+.11**
H_{19}	Fit between parent brand and extension → retailer acceptance	—	+.23**	+.23**	+.23**
H_{20}	Perceived risk of unknown brands → consumer innovativeness	—	–.25**	–.25**	–.25**
H_{21}	Fit × quality of the parent brand → extension success	—	—	+.07**	—
H_{24}	Fit × parent-brand conviction → extension success	—	—	—	+.08**

History × fit; marketing support × retailer acceptance; perceived risk of unknown brands × quality; history × quality: n.s.

Total Effects of the Success Factors (i.e., Significant Direct Effect + Significant Indirect Effects) [Percentage Relative Importance in Brackets]

	M1	M2	M3a	M3b
Quality of the parent brand	.10 [7.19]	.11 [4.76]	.11 [4.62]	.11 [4.60]
History of previous brand extensions	–.04 [2.88]	.09 [3.90]	.09 [3.78]	.09 [3.77]
Parent-brand conviction	.25 [17.99]	.38 [16.45]	.38 [15.97]	.38 [15.90]
Parent-brand experience	n.s. [.00]	.24 [10.39]	.24 [10.08]	.24 [10.04]
Marketing support	.13 [9.35]	.52 [22.51]	.52 [21.85]	.52 [21.76]
Retailer acceptance	.29 [20.86]	.30 [12.99]	.30 [12.61]	.30 [12.55]
Fit	.45 [32.37]	.53 [22.94]	.54 [22.69]	.54 [22.59]
Perceived risk of unknown brands	.08 [5.76]	.08 [3.46]	.07 [2.94]	.06 [2.51]
Consumer innovativeness	.05 [3.60]	.06 [2.60]	.06 [2.52]	.07 [2.39]
Linkage of the utility of the parent brand to product attributes	n.s. [.00]	n.s. [.00]	n.s. [.00]	n.s. [.00]
Fit × quality of the parent brand	—	—	.07 [2.94]	—
Fit × parent-brand conviction	—	—	—	.08 [3.35]
NFI	.93	.92	.92	.92
NNFI	.93	.92	.92	.92
CFI	.95	.93	.93	.93
RMSEA	.045	.048	.046	.046
p value of test of close fit	1.000	.995	1.000	1.000
n (calibration sample)	3068	3068	3068	3068

*Significant at the $p < .05$ level

**Significant at the $p < .01$ level

Notes: n.s. = not significant. NFI = normed fit index, NNFI = nonnormed fit index, CFI = comparative fit index, and RMSEA = root mean square error of approximation.

of previous brand extensions on extension success is significant but negative. Thus, H_2 is not supported.[5] However, Models M2 and M3 reveal that this factor is a significant, positive determinant of brand extension success as a result of its significant and positive effects on parent-brand conviction, quality of the parent brand, and retailer acceptance.

---

[5] The negative direct influence might be attributed to some kind of parent-brand dilution effect. A history of brand extensions may cause increasing incongruity in the parent brand's product range. A review of the data indicates that respondents assessed nearly one-half of the selected parent brands as having only a moderately clear brand image (seven-point rating scale: –3 = "very unclear," and +3 = "very clear"; for ten brands, means range from –1 to .5).

Effects of the extension's marketing context. All models support the direct effect on brand extension success associated with marketing support of the extension. However, Model M1 again provides a biased estimate of the impact of this success factor on extension success. Models M2 and M3 reveal that marketing support has both a direct impact on extension success and three indirect effects on extension success through the following paths:

1. Marketing support → fit → extension success.
2. Marketing support → fit → retailer acceptance → extension success.
3. Marketing support → retailer acceptance → extension success.

Accordingly, marketing support has a substantial total effect of .52 on the success of a brand extension. Consequently, a considerable change emerges in the relative influence of this variable compared with Model M1, which ignores structural relationships among the success factors. The relative influence of marketing support on the success of a brand extension increases from 9.35% (M1) to 22.51% (M2). In general, the total effect of any variable X_1 on any variable X_2 represents the comprehensive impact of X_1 on X_2, when the entire system of structural relationships is considered among the variables under investigation. As with the total effect of the marketing support variable, the total effects of all other success factors can be computed to determine their relative importance for the success of a brand extension. Similarly, all models find strong support for the significance of the effect related to retailer acceptance. However, Table 3 shows that M1 overestimates the relevance of this factor for the success of an extension product.

Effects of the relationship between the parent brand and the extension product. The linkage of the utility of the parent brand to specific product attributes is statistically insignificant in all models, whereas the fit between the parent and the extension category emerges as the most important factor (sign as expected). Although the significance of this factor was not unexpected (given the results of prior research; e.g., Bottomley and Holden 2001), the overwhelming impact of this factor compared with other brand extension success factors is a notable result, especially because we used real brand extensions. Not too long ago, researchers repeatedly presumed that the overwhelming impact of the fit could be attributed to the use of hypothetical brand extensions (e.g., Klink and Smith 2001). However, the results of our study show that the fit still has a substantial effect on brand extension success, even under exclusive use of real brand extensions. Note that we applied a widely accepted and comprehensive measure of perceived fit (see the Appendix). The total effects reveal a considerable change in the relative influence of perceived fit on brand extension success compared with Model M1, which ignores structural relationships among the success factors. Relative importance of the fit between the parent brand and the extension product decreases from 32.4% to 22.6% (though fit still remains the most important success factor).

Effects of the extension's product category characteristics. Perceived risk of unknown brands and consumer innovativeness emerged as significant determinants of extension success in all models. However, these factors exhibited only minor relevance.

Moderator effects. We applied the single-indicator technique to incorporate the effects of moderating variables (Ping 1995, 1998). We found statistically significant effects for only two interaction terms (see Table 3), namely, the interaction terms of fit with the quality of the parent brand and with the parent-brand conviction. However, the contribution of the two interaction terms was small compared with the main effects of the success factors (direct and indirect effects of the variables). This means that moderating effects play a relatively minor role and that the incremental variance explained by the interaction terms is small. Simultaneous estimation of both interaction effects supports their statistical significance. However, the various fit indicators are significantly lower than for M2, and they suggest that the model does not represent an acceptable fit to the data (normed fit index = .87, nonnormed fit index = .87, comparative fit index = .88, and root mean square error of approximation = .058; p value of the test of close fit = .000).

Finally, the degree to which a consumer is familiar with the extension might be an important moderating variable because we used real brand extensions. For example, parent-brand characteristics (see Table 1) and the fit between the parent brand and the extension product should have a much greater effect when the extension is relatively unfamiliar. We applied a subgroup analysis (e.g., Jöreskog 1971; Ping 1998) to test the moderating effect of consumers' familiarity. The procedure involved dividing our sample into two groups of cases on the basis of different levels of familiarity (low and high) and estimating M2 in each subgroup without coefficient equality constraints. We measured consumers' familiarity with the extension by asking respondents how often they used the extension product (–3 = "not at all," and +3 = "very frequently"). We divided our sample into two groups on the basis of a mean split. The low- (high-) familiarity subgroup comprised respondents below (above) the mean. Table 4 reports the parameter estimates and goodness-of-fit indicators of the two-group model (M4).

In many ways, the results of the two-group analysis support the analysis that is based on the total sample. The estimates are quite robust. We find notable differences between the two groups for only three variables. The quality of the parent brand and the marketing support of the extension have a greater impact when the extension is relatively unfamiliar. This seems reasonable because respondents, who are unfamiliar with the extension product, will rely more heavily on external cues (e.g., brand name, marketing support) when evaluating the extension. The fit between the parent brand and the extension product has a greater effect when consumers are relatively familiar with the extension. This result is notable. We expected that the fit would have a greater impact within the group of consumers who were relatively unfamiliar with the extension product. However, high familiarity might help consumers identify or create

TABLE 4
Moderating Effect of Consumers' Brand Extension Familiarity (M4)

Paths	Estimates Low Familiarity (Below Mean)	 High Familiarity (Above Mean)	 Total Sample
H_1 Quality of the parent brand → extension success	+.20**	+.11**	+.11**
H_2 History of previous brand extensions → extension success	n.s.	n.s.	n.s.
H_3 Parent-brand conviction → extension success	+.31**	+.31**	+.30**
H_4 Parent-brand experience → extension success	n.s.	n.s.	n.s.
H_5 Marketing support → extension success	+.21**	+.05*	+.14**
H_6 Retailer acceptance → extension success	+.25**	+.26**	+.30**
H_7 Fit → extension success	+.46**	+.53**	+.46**
H_8 Linkage of the utility of the parent brand to specific product attributes → extension success	n.s.	n.s.	n.s.
H_9 Perceived risk of unknown brands → extension success	+.08**	+.09**	+.09**
H_{10} Consumer innovativeness → extension success	+.09**	+.06*	+.06**
H_{11} Quality of the parent brand → retailer acceptance	n.s.	n.s.	n.s.
H_{12} Marketing support → retailer acceptance	+.56**	+.39**	+.56**
H_{13} History of previous brand extensions → parent-brand conviction	+.14**	+.16**	.13**
H_{14} Parent-brand experience → parent-brand conviction	+.60**	+.61**	.64**
H_{15} Marketing support → fit	+.30**	+.32**	+.40**
H_{16} Parent-brand conviction → quality of the parent brand	+.73**	+.72**	+.74**
H_{17} History of previous brand extensions → quality of the parent brand	+.05*	+.08**	+.05*
H_{18} History of previous brand extensions → retailer acceptance	+.12**	+.18**	+.11**
H_{19} Fit between parent brand and extension → retailer acceptance	+.19**	+.29**	+.23**
H_{20} Perceived risk of unknown brands → consumer innovativeness	−.26**	−.25**	−.25**

Total Effects of the Success Factors (i.e., Significant Direct Effect + Significant Indirect Effects) [Percentage Relative Importance in Brackets]

Quality of the parent brand	.20 [8.20]	.11 [4.98]	.11 [4.76]
History of previous brand extensions	.10 [4.10]	.11 [4.98]	.09 [3.90]
Parent-brand conviction	.46 [19.74]	.39 [17.65]	.38 [16.45]
Parent-brand experience	.27 [11.07]	.24 [10.86]	.24 [10.39]
Marketing support	.50 [20.49]	.35 [15.84]	.52 [22.51]
Retailer acceptance	.25 [10.25]	.26 [11.76]	.30 [12.99]
Fit	.51 [20.90]	.61 [27.60]	.53 [22.94]
Perceived risk of unknown brands	.06 [2.46]	.08 [3.61]	.08 [3.46]
Consumer innovativeness	.09 [3.69]	.06 [2.71]	.06 [2.60]
Linkage of the utility of the parent brand to specific product attributes	n.s.	n.s.	n.s.

	Two-Groups Model	Total Sample
NFI	.91	.92
NNFI	.91	.92
CFI	.92	.93
RMSEA	.033	.048
p value of test of close fit	1.000	.995

*Significant at the $p < .05$ level.
**Significant at the $p < .01$ level.
Notes: n = 3853 (below mean)/2275 (above mean). n.s. = not significant. NFI = normed fit index, NNFI = nonnormed fit index, CFI = comparative fit index, and RMSEA = root mean square error of approximation.

more shared attributes between the extension and the brand schema (e.g., Klink and Smith 2001). As a result, higher levels of familiarity may increase the impact of fit on consumers' extension evaluations.

To determine whether the differences between the two groups were significant, we constrained the model coefficients to be equal between subgroups, and we reestimated the model. We compared the resulting chi-square statistic with the chi-square statistic from the first estimation. We found a statistically significant chi-square difference between the two groups, which suggests that the differences are significant.

Model comparison. We compared the performance of the estimated models using nested and nonnested model comparisons (Leeflang et al. 2000) based on the results of the cross-validation. Model M2 emerged as the "best" model. The chi-square difference for M2 was substantially lower than that for the other models; that is, applying the constrained model (constrained path coefficients, factor loadings, and variance–covariance of the constructs) had the least negative consequence for Model M2. Mere addition of parameters should improve model fit. Therefore, it was important to consider the degrees of freedom when we compared the performance of the alternative models. The chi-square difference divided by the difference in the degrees of freedom was lower for M2 than for the other models. Therefore, incorporating indirect effects into M1 (resulting in M2) substantially improved prediction, whereas the incremental improvement by the interaction terms, which we introduced after accounting for the indirect effects of the ten success factors, was small.

Empirical Application

To illustrate the managerial relevance of our findings, we applied our best-fitting Model M2 to several hypothetical extension decisions in an additional empirical study and compared it with an intuitive model that managers who are not aware of M2 might use. We knew from our first expert survey that, in general, nearly all 21 managers agreed that the ten success factors implemented in M2 were relevant for brand extension success. However, the experts were not able to agree on the relative importance weights for the ten success factors. Thus, an intuitive model that managers might use would assume equal importance weights for all success factors (i.e., M1 with equal importance weights). Indeed, many managers with whom we spoke seemed to use some kind of a heuristic technique with equally weighted success factors to arrive at expected performance judgments.

In our application, we considered a manager who needed to decide whether to introduce several new products by extending established brands. We focused our illustrative cases on 12 typical, realistic brand extensions that had not been introduced yet (e.g., aspirin, adhesive bandages, Mon Cheri ice cream). The 12 brand extensions stemmed from six well-known FMCG brands in Germany. To apply M2 in this context, we needed data from a consumer survey on the levels of the ten brand extension success factors (seven-point rating scales) under consideration, in addition to the estimated parameters in Table 3. We administered a corresponding questionnaire to a convenience sample of 119 marketing students to obtain the required data. The application of M2 to the collected data provided estimates of the expected overall extension success and the incremental influence of each driver of brand extension success. This kind of information helps managers decide whether to introduce a proposed extension.

On the basis of the collected data, we assessed expected performance of the hypothetical extension products by taking (1) the average of the ten success factors (i.e., M1 with equal importance weights) and (2) the weighted average of the ten success factors using the empirically determined importance weights (M2). We scaled the importance weights so that they added up to 100%. Therefore, an extension product with the lowest rating on all success factors earns an expected performance score of 1, and an extension product with the highest rating on each factor earns a score of 7. We found notable differences between M2 and the intuitive model that managers advocated. Differences in expected performance scores achieved up to .5 points on a seven-point rating scale, which can be viewed as substantial.

Furthermore, the assumption of equal weights for the success factors resulted in faulty interpretation of the incremental effect of the success factors on the expected performance of an extension product. The estimated path coefficients of the various success factors are a possible measure for the incremental effect of a certain success factor on brand extension success. However, elasticity is the usual measure for the incremental effect of a variable. Therefore, using (1) a model with equal weights of the success factors and (2) the empirically determined importance weights (M2), we calculated the resulting increase in expected performance (expressed as a percentage) if a certain success factor is increased by 10%. The resulting elasticity measured the incremental effect on expected performance (expressed as a percentage) of a 10% increase in a specified success factor. Table 5 reports the calculated elasticities for all success factors averaged across the extension products. There are substantial differences between the two models for nearly all elasticities. For example, a 10% increase in marketing support results in a 2.52% increase of expected performance in Model M2 versus a 1.14% increase in a model with equal weights. This means that the assumption of equal weights substantially underestimates the effect of marketing support on expected performance. The incremental effects that are associated with certain variables are the key factors in the decision of how to allocate resources to them. Thus, an intuitive model with equal weights can result in largely inappropriate allocation of resources.

Summary

We tested several conceptual models that considered brand extensions within a comprehensive nomological net by building on the results of prior research and managerial wisdom derived from expert judgments. Our simultaneous analysis of a broad variety of success factors revealed nine factors that significantly influence brand extension success. Fit between the parent brand and the extension product, marketing support, parent-brand conviction, retailer acceptance, and parent-brand experience were particularly major contributors in driving brand extension success. These factors deserve managerial attention. In addition, we found several important structural relationships among the investigated success factors. An analysis that fails to account for these relationships causes faulty interpretation of the significance, the relative importance, and the incremental influence of success factors. In addition, we found statistically significant effects for two interaction terms: the interaction of fit with the quality of the parent brand and with parent-brand conviction. However, the contribution of the two interaction terms is small compared with the main effects of the success factors. We illustrated the managerial relevance of our findings in an additional empirical study.

TABLE 5
Comparing M2 with an Intuitive Model (M1 with Equal Weights)

	Model M2		Equal Weights		
Success Factor	**(a) Elasticity[a]**	**(b) SD**	**(c) Elasticity[a]**	**(d) SD**	**(a)/(c)**
Quality of the parent brand	.52	.18	1.10	.33	.47
History of previous brand extensions	.31	.08	.82	.23	.38
Parent-brand conviction	1.64	.59	1.01	.31	1.62
Parent-brand experience	.99	.18	.98	.13	1.01
Marketing support	2.52	.59	1.14	.21	2.20
Retailer acceptance	1.26	.35	1.02	.33	1.24
Fit	2.26	1.23	1.05	.62	2.15
Linkage of the utility of the parent brand to specific product attributes	.00	.00	1.14	.19	.00
Perceived risk of unknown brands	.27	.04	.81	.13	.34
Consumer innovativeness	.24	.03	.94	.09	.25

[a] Incremental effect (percentage) on expected performance of a 10% increase in the success factor. For example, an increase in marketing support of 10% results in a 2.52% increase of expected performance in Model M2 versus a 1.14% increase in a model with equal weights (i.e., equal weights substantially underestimate the effect of marketing support on expected performance).

Discussion

Managerial Implications

Success factors that managers should care about. From a managerial perspective, it is especially important to know the relevant determinants of extension success. Our analyses help managers divide the large number of potentially relevant success factors into essential factors (i.e., fit, marketing support, retailer acceptance, parent-brand conviction, and parent-brand experience) and less relevant or unimportant factors (i.e., history of previous extensions, consumer innovativeness, linkage of the utility of the parent brand to specific product attributes, and moderating effects). Managers should pay the greatest attention to the five success factors that substantially influence brand extension success. Consequently, the question arises of how to influence these factors to increase the likelihood of extension success. Our research uncovers several important mediating effects that may serve as a starting point for the identification of promising ways to influence certain success determinants and, consequently, the likelihood of extension success. The following points focus on the issue of how managers can influence the five most important success factors identified in our study:

- *Fit between parent brand and extension product*: Appropriate selection of the parent brand and extension product category can directly affect fit. Extension advertisements can also increase the salience of crucial brand associations that help consumers infer extension features and benefits and thus understand how an extension fits. Consumers will likely infer judgments of fit and consistency when an advertisement illustrates how parent-brand attributes improve the extension's ability to provide benefits. Repeated exposure to advertisements that evoke appropriate parent-brand associations also helps consumers establish linkages between the brand and the extension product, thus positively influencing perceived fit between the extension product and the extending brand (Lane 2000).
- *Parent-brand conviction and parent-brand experience*: Although extension products do not guarantee success on the basis of the brand name alone, our results indicate that parent-brand characteristics, such as consumers' parent-brand experience and conviction, play an important role in driving brand extension success. Managers frequently cannot influence these two factors in the short or medium term, because both factors reflect specific characteristics of the parent brand. However, our results indicate that building customer-based brand equity and/or acquiring strong brands are favorable strategies because they represent a prerequisite for the successful leveraging of an existing brand to reap some financial benefit. Moreover, managers who have access to a portfolio of parent brands might consider influencing parent-brand conviction and parent-brand experience by selecting an appropriate brand that performs favorably in terms of these factors.
- *Retailer acceptance*: There are various possible ways to influence retailers' acceptance of a new product. For example, previous research has shown the positive effects of trade and slotting allowances on retailer acceptance of new products (e.g., Collins-Dodd and Louviere 1999). Promotional allowances reduce retailers' costs of telling consumers that the extension product can be obtained in a given store, thus strengthening interstore competition. Consumer advertising also builds demand and therefore should have a positive effect on retailers' decisions to accept new listings.
- *Marketing support*: The marketing support that the extension product receives plays a critical role for FMCGs in determining the success of the new product. This factor is of specific interest for managers because it is under a company's direct control and, in general, can be influenced in the short term. However, the financial well-being of the company represents a boundary condition regarding the total support that new introductions may receive.

Relative importance of success factors. Another set of implications for managers is related to the incremental influence of each success factor on the performance of the extension product. Thus far, little has been known about the relative importance of success factors in explaining brand extension success. Therefore, managers might intuitively assume equal importance weights for all success factors. Our results indicate that relative importance varies substantially across the success factors. To illustrate the managerial relevance of this finding, we applied the model advocated by managers and our best model (M2) to 12 hypothetical extension decisions. Our illustrative cases show that the assumption of equal weights for the success factors results in faulty prediction of extension success and in faulty interpretation of the incremental effect of the success factors on expected performance of an extension product. This is a serious problem because it is imperative for managers to know the magnitude of effects so that they can analyze the reasons for brand extension failure/success and properly allocate resources to the various success factors that are under a firm's control. Managers need to be aware that simple heuristic techniques, which ignore the differences in importance weights among the success factors, are likely to cause notable deviations from marketplace performance. Our results demonstrate that managers need to pay attention to the varying incremental effects of success determinants when evaluating and managing extension products.

Limitations and Further Research

Measurement of the relative importance of brand extension success factors is a complex issue. It has been one of the key issues in prior research, but this work has provided limited insights into the relative importance of success factors because it has only analyzed a small fraction of success factors at one time. Against this background, our study is more appropriate for understanding relative importance. It synthesizes prior work and simultaneously examines the direct and indirect effects of a large number of potentially relevant success factors. However, note that the range and variance of brand extension success factors, among others, drive relative importance. We selected a broad variety of parent brands and extension products to ensure that the brand extensions used in our study varied with respect to consumers' judgments of success factors. The selected parent brands belong to major firms on the German FMCG market. The extension products covered the ten most important food and nonfood product categories of ACNielsen's well-known classification. The manipulation check of success factors confirmed a large empirical variance. In this context, we note that the majority of previous studies also did not vary consumers' judgments of the success factors by means of an experimental design. A notable exception might be the fit variable that was often manipulated in previous work by the selection of low-fit and high-fit extensions. We realized the latter in our study and thus selected parent brands and extension products appropriately. Nevertheless, it might be useful for validation purposes if future work were to replicate our study in a laboratory setting using hypothetical parent brands and extension products that enable explicit manipulation of success factors. However, use of hypothetical parent brands represents a serious threat to the external validity of the results (e.g., Boush and Loken 1991).

Furthermore, we chose to analyze only FMCGs. It would be interesting to investigate the extent to which our findings could be generalized to other fields, such as consumer durables or services. Further research is also needed to investigate the generalizability of our results across multiple categories in FMCG.

In conclusion, fit should become a focus of further research, given the importance of the fit variable found in our study and in previous studies (e.g., Bottomley and Holden 2001). There is no universally accepted conceptualization of fit, which raises the more general question, What constitutes fit? The answer to this question is especially important for understanding how managers can manipulate perceived fit.

APPENDIX
Measures

Success Factor/Item	Source
Quality of the Parent Brand	
Perceived overall quality of the flagship product (–3 = "inferior," and 3 = "superior")	Aaker and Keller (1990); Sheinin and Schmitt (1994)
Extent to which participants agreed/disagreed with the following statements (–3 = "strongly disagree," and 3 = "strongly agree"):	
[Brand name] offers high-quality products.	
The quality of [brand name] products is far above average.	
History of Previous Brand Extensions	
Number of products affiliated with the brand (–3 = "very few products," and 3 = "a lot of products")	Dacin and Smith (1994); Keller and Aaker (1992a); Sheinin and Schmitt (1994)
Success of the products affiliated with the brand (–3 = "not at all successful," and 3 = "very successful")	
Degree of similarity between the products affiliated with the brand (–3 = "not very similar," and 3 = "very similar")	
Participants were instructed to answer questions about products previously introduced and not the new extension product. This clarified that the question referred to all products previously introduced under the brand name and not to the new extension product.	

APPENDIX
Continued

Success Factor/Item	Source
Parent-Brand Conviction	
Extent to which participants agreed/disagreed with the following statements (–3 = “strongly disagree,” and 3 = “strongly agree”): In evaluating a new [product category] product, I could trust [brand name]. [Brand name] is a likeable brand. I relate to [brand name].	DelVecchio (2000); Kirmani, Sood, and Bridges (1999)
Parent-Brand Experience	
Frequency of using the parent brand (–3 = “not very often,” and 3 = “very often”) Frequency of purchasing the parent brand (–3 = “not very often,” and 3 = “very often”) Intention to buy the parent brand in the future (–3 = “not very likely,” and 3 = “very likely”)	Broniarczyk and Alba (1994); Swaminathan, Fox, and Reddy (2001)
Marketing Support	
Extent to which participants agreed/disagreed with the following statements (–3 = “strongly disagree,” and 3 = “strongly agree”): [Extension product] is well supported in terms of advertising. [Extension product] receives competent marketing support.	Nijssen (1999); Reddy, Holak, and Bhat (1994)
Retailer Acceptance	
Extent to which participants agreed/disagreed with the following statements (–3 = “strongly disagree,” and 3 = “strongly agree”): [Extension product] is well supported in terms of distribution. [Extension product] is available in many supermarkets.	Nijssen (1999)
Fit	
Global similarity between the parent brand and the extension product (–3 = “not very similar,” and 3 = “very similar”) Would the people, facilities, and skills used in making the original product be helpful if the manufacturer were to make the extension product? (–3 = “not at all helpful,” and 3 = “very helpful”) Extent to which parent-brand-specific associations are relevant in the extension category: Step 1: stating of brand associations; Step 2: relevance of these associations in the extension category (– 3 = “not at all relevant,” and 3 = “very relevant”).	Aaker and Keller (1990); Bottomley and Doyle (1996); Broniarczyk and Alba (1994)
Linkage of the Utility of the Parent Brand to Product Attributes of the Original Product Category	
Extent to which participants agreed/disagreed with the following statement (–3 = “strongly disagree,” and 3 = “strongly agree”): [Brand name] is closely tied to the attributes of the original product category. My associations with [brand name] are closely tied to the attributes of the original product category.	Rangaswamy, Burke, and Oliva (1993)
Perceived Risk of Unknown Brands	
Extent to which participants agreed/disagreed with the following statements (–3 = “strongly disagree,” and 3 = “strongly agree”): If I buy an unknown brand, I would feel very uncertain of the level of quality that I am getting. I prefer buying a well-known brand, because I need the reassurance of an established brand name. I prefer buying a well-known brand, because the risk that my needs will not be met is low compared with an unknown brand.	DelVecchio (2000)
Innovativeness	
Extent to which participants agreed/disagreed with the following statements (–3 = “strongly disagree,” and 3 = “strongly agree”): Overall, I enjoy buying the latest products. I like to purchase new products before others do. Overall, it is exciting to buy the latest products.	Hem, Chernatony, and Iversen (2001); Klink and Smith (2001); Midgley and Dowling (1978)

REFERENCES

Aaker, David A. and Kevin Lane Keller (1990), "Consumer Evaluations of Brand Extensions," *Journal of Marketing*, 54 (January), 27–41.

Arbuckle, Jim L. and Werner Wothke (1999), *Amos 4.0 User's Guide*. Chicago: SmallWaters Co.

Bagozzi, Richard P. and Youjae Yi (1988), "On the Evaluation of Structural Equation Models," *Journal of the Academy of Marketing Science*, 16 (1), 74–94.

Barone, Michael J., Paul W. Miniard, and Jean B. Romeo (2000), "The Influence of Positive Mood on Brand Extension Evaluation," *Journal of Consumer Research*, 26 (March), 386–400.

Baumgartner, Hans and Christian Homburg (1996), "Applications of Structural Equation Modeling in Marketing and Consumer Research: A Review," *International Journal of Research in Marketing*, 13 (2), 139–61.

Bollen, Kenneth A. (1989), *Structural Equations with Latent Variables*. New York: John Wiley & Sons.

Bottomley, Paul A. and John R. Doyle (1996), "The Formation of Attitudes Towards Brand Extensions: Testing and Generalising Aaker and Keller's Model," *International Journal of Research in Marketing*, 13 (4), 365–77.

———and Stephen J.S. Holden (2001), "Do We Really Know How Consumers Evaluate Brand Extensions? Empirical Generalizations Based on Secondary Analysis of Eight Studies," *Journal of Marketing Research*, 38 (November), 494–500.

Boush, David M. and Barbara Loken (1991), "A Process-Tracing Study of Brand Extension Evaluation," *Journal of Marketing Research*, 28 (February), 16–28.

Broniarczyk, Susan M. and Joseph W. Alba (1994), "The Importance of the Brand in Brand Extension," *Journal of Marketing Research*, 31 (May), 214–28.

Byrne, Barbara M. (2001), *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. Mahwah, NJ: Lawrence Erlbaum Associates.

Collins-Dodd, Colleen and Jordan J. Louviere (1999), "Brand Equity and Retailer Acceptance of Brand Extensions," *Journal of Retailing and Consumer Services*, 6 (1), 1–13.

Dacin, Peter A. and Daniel C. Smith (1994), "The Effect of Brand Portfolio Characteristics on Consumer Evaluations of Brand Extension," *Journal of Marketing Research*, 31 (May), 229–42.

Dawar, Niraj and Paul F. Anderson (1994), "The Effects of Order and Direction on Multiple Brand Extensions," *Journal of Business Research*, 30 (2), 119–29.

DelVecchio, Devon (2000), "Moving Beyond Fit: The Role of Brand Portfolio Characteristics in Consumer Evaluations of Brand Reliability," *Journal of Product & Brand Management*, 9 (7), 457–71.

Desai, Kalpesh Kaushik and Kevin L. Keller (2002), "The Effects of Ingredient Branding Strategies on Host Brand Extendibility," *Journal of Marketing*, 66 (January), 73–93.

Edison, Steve W. and Gary L. Geissler (2003), "Measuring Attitudes Towards General Technology: Antecedents, Hypotheses, and Scale Development," *Journal of Targeting, Measurement and Analysis for Marketing*, 12 (2), 137–56.

Ernst & Young and ACNielsen (1999), *New Product Introduction: Successful Innovation/Failure: A Fragile Boundary*. Paris: Ernst & Young Global Client Consulting.

Ford, Barry L. (1976), "Missing Data Procedures. A Comparative Study," in *ASA Proceedings of the Social Statistics Section*. Washington, DC: American Statistical Association, 324–29.

Fornell, Claes and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39–50.

Gustafsson Jan-Eric and Per Arne Stahl (2000), *STREAMS User's Guide Version 2.5 for Windows*. Mölndal, Sweden: MultivariateWare.

Hem, Leif E., Leslie de Chernatony, and Nina M. Iversen (2001), "Factors Influencing Successful Brand Extensions," working paper, Birmingham University Business School, Birmingham, UK.

Hu, Li-tze and Peter M. Bentler (1995), "Evaluating Model Fit," in *Structural Equation Modeling: Concepts, Issues, and Applications*, Rick H. Hoyle, ed. Thousand Oaks, CA: Sage Publications, 76–99.

Jöreskog, Karl G. (1971), "Simultaneous Factor Analysis in Several Populations," *Psychometrika*, 36 (4), 409–426.

Keller, Kevin Lane and David A. Aaker (1992a), "The Effects of Corporate Images and Branding Strategies on New Product Evaluations," Working Paper No. 1216, Graduate School of Business, Stanford University.

——— and ——— (1992b), "The Effects of Sequential Introduction of Brand Extensions," *Journal of Marketing Research*, 29 (February), 35–50.

Kirmani, Amna, Sanjay Sood, and Sheri Bridges (1999), "The Ownership Effect in Consumer Responses to Brand Line Stretches," *Journal of Marketing*, 63 (January), 88–101.

Klink, Richard R. and Daniel C. Smith (2001), "Threats to the External Validity of Brand Extension Research," *Journal of Marketing Research*, 38 (August), 326–35.

Lane, Vicki R. (2000), "The Impact of Ad Repetition and Ad Content on Consumer Perceptions of Incongruent Extensions," *Journal of Marketing*, 64 (April), 80–91.

——— and Robert Jacobson (1995), "Stock Market Reactions to Brand Extension Announcements: The Effects of Brand Attitude and Familiarity," *Journal of Marketing*, 59 (January), 63–77.

Laroche, Michel, Jasmin Bergeron, and Christine Goutaland (2001), "A Three-Dimensional Scale of Intangibility," *Journal of Service Research*, 4 (1), 26–38.

Leeflang, Peter S.H and Dick R. Wittink (2000), "Building Models for Marketing Decisions: Past, Present, and Future," *International Journal of Research in Marketing*, 17 (2–3), 105–26.

———, ———, Michel Wedel, and Philippe A. Naert (2000), *Building Models for Marketing Decisions*. Boston: Kluwer Academic Publisher.

Marketing (UK) (2003), "Premium Extensions Are Proving to Be the Most Promising FMCG Launches, as Manufacturers Look to Counteract Retailers' Price Cuts," (August 28), 25.

Midgley, David F. and Grahame R. Dowling (1978), "Innovativeness: The Concept and Its Measurement," *Journal of Consumer Research*, 4 (March), 229–42.

Mullen, Michael R., George R. Milne, and Patricia Doney (1995), "An International Marketing Application of Outlier Analysis for Structural Equations: A Methodological Note," *Journal of International Marketing*, 3 (1), 45–62

Muthén, Bengt (1989), "Latent Variable Modeling in Heterogenous Populations," *Psychometrika*, 54 (4), 557–85

——— (1990), "Mean and Covariance Structure Analysis of Hierarchical Data," *UCLA Statistics Series*, Vol. 62. Los Angeles: University of California.

Nijssen, Edwin J. (1999), "Success Factors of Line Extensions of Fast-Moving Consumer Goods," *European Journal of Marketing*, 33 (5–6), 450–69.

——— and Pete Bucklin (1998), "The Effect of Involvement upon Brand Extension," working paper, Nijmegen School of Management, University of Nijmegen.

Ping, Robert A. (1995), "A Parsimonious Estimating Technique for Interaction and Quadratic Latent Variables," *Journal of Marketing Research*, 32 (August), 336–47.

——— (1998), "EQS and LISREL Examples Using Survey Data," in *Interaction and Nonlinear Effects in Structural Equation Modeling*, R.E. Schumacker and G.A. Marcoulides, eds. Mahwah, NJ: Lawrence Erlbaum Associates, 63–100.

Rangaswamy, Arvind, Raymond R. Burke, and Terence A. Oliva (1993), "Brand Equity and the Extendibility of Brand Names," *International Journal of Research in Marketing*, 10 (1), 61–75.

Reddy, Srinivas K., Susan L. Holak, and Subodh Bhat (1994), "To Extend or Not to Extend: Success Determinants of Line Extensions," *Journal of Marketing Research*, 31 (May), 243–62.

Rogers, Everett M. (1983), *Diffusion of Innovations*, 3d ed. New York: The Free Press.

Sandor, Zsolt and Michel Wedel (2001), "Designing Conjoint Choice Experiments Using Managers' Prior Beliefs," *Journal of Marketing Research*, 38 (4), 430–43.

Sheinin, Daniel A. and Bernd H. Schmitt (1994), "Extending Brands with New Product Concepts: The Role of Category Attribute Congruity, Brand Affect, and Brand Breadth," *Journal of Business Research*, 31 (1), 1–10.

Smith, Daniel C. and C. Whan Park (1992), "The Effects of Brand Extensions on Market Share and Advertising Efficiency," *Journal of Marketing Research*, 29 (August), 296–313.

Steenkamp, Jan-Benedict E.M. and Hans C.M. van Trijp (1991), "The Use of LISREL in Validating Marketing Constructs," *International Journal of Research in Marketing*, 8 (4), 283–99.

Sunde, Lorraine and Roderick J. Brodie (1993), "Consumer Evaluations of Brand Extensions: Further Empirical Results," *International Journal of Research in Marketing*, 10 (1), 47–53.

Swaminathan, Vanitha, Richard J. Fox, and Srinivas K. Reddy (2001), "The Impact of Brand Extension Introduction on Choice," *Journal of Marketing*, 65 (October), 1–15.

Tauber, Edward M. (1988), "Brand Leverage: Strategy for Growth in a Cost-Control World," *Journal of Advertising Research*, 28 (August–September), 26–30.

Sonya A. Grier, Anne M. Brumbaugh, & Corliss G. Thornton

Crossover Dreams: Consumer Responses to Ethnic-Oriented Products

The authors develop and evaluate a framework for investigating and understanding ethnic product crossover, that is, when a product intended for one ethnic minority group gains significant penetration among consumers outside the referent ethnic group. In three studies, the authors investigate how a product's characteristics, the promotion and distribution decisions made for the product, and consumers' propensity for diversity influence the product's likelihood of crossing over from the intended ethnic target market to mainstream white consumers. Product characteristics interact with both other marketing decisions and consumers' diversity-seeking tendencies to influence whether consumers will be interested in ethnic products and the social context in which they are willing to consume them. The authors discuss the implications of the findings for theory and practice and provide directions for further research that include consideration of the product's ethnic embeddedness, the context in which the product will be consumed, and consumers' diversity-seeking tendencies.

Marketers of ethnic-oriented niche products frequently strategize to create crossover of their products into larger markets for economies of scale and commercial viability. However, when such products are closely identified with the specialized target market they serve, successful crossover to a mainstream market can be a challenge. Nevertheless, some formerly ethnic-oriented niche products successfully cross over to the mainstream market. For example, 60% of hip-hop and rap, traditionally black music, is purchased primarily by white youths (Hughes 2002); salsa surpassed ketchup as the top-selling condiment in the United States in 1998 (Mellgren 2001); 90% of sushi chef students at a Los Angeles academy are not Japanese; and more than 1000 U.S. grocery stores have sushi counters (Walkup 2003). In contrast, however, some ethnic-oriented niche products do not cross over despite marketers' intentions for them to do so. For example, annual lists of top-ten television programs for black and white consumers demonstrate strongly segregated audiences: Blacks watch shows targeted to a black audience, and whites watch shows intended for a white or mainstream audience (Gibbs 1999). Despite script and cast changes in response to marketing research, producers could not induce significant crossover among Anglo moviegoers for *Price of Glory*, a *Rocky*-like movie about a Hispanic fighter (Bannon 2000). Why do some ethnic-oriented products cross over but others do not?

We examine the crossover of ethnic-oriented products into the larger mainstream market. Specifically, we address three key research questions:

1. How do an ethnic-oriented product's characteristics influence its appeal among consumers of the referent ethnic minority and the mainstream ethnic majority?
2. How do other marketing activities influence the appeal of and intended behaviors toward ethnic-oriented products among mainstream ethnic majority consumers?
3. How do consumers differ in their willingness to purchase and consume ethnic-oriented products?

We address these questions by first developing a framework for understanding ethnic-oriented product crossover. Then, in three studies, we examine several aspects of this framework, including how ethnic congruence, familiarity, diversity orientation, and social context influence consumers' attitudes and purchase intentions toward ethnic-oriented products. We discuss the implications of our findings in relation to our framework and describe how our results inform our practical understanding of when crossover is and is not likely to occur. We also discuss how the framework may be applied to the understanding of crossover potential across product categories, consumption situations, and other social identity factors (e.g., age, gender).

A Framework for Ethnic Crossover

Crossover occurs when a product designed for one target segment meets with attraction and acceptance in another segment (Garafalo 1993; Gibbs 1999). Crossover has traditionally been defined as the movement of a product from margin to mainstream, but not vice versa (Garafalo 1994). We focus on ethnic crossover, whereby a product that

---

Sonya A. Grier is Robert Wood Johnson Health & Society Scholar, University of Pennsylvania (e-mail: griers@wharton.upenn.edu). Anne M. Brumbaugh is Assistant Professor of Marketing, Babcock School of Management, Wake Forest University (e-mail: anne.brumbaugh@mba.wfu.edu). Corliss G. Thornton is Associate Professor of Marketing, Robinson College of Business, Georgia State University (e-mail: cthornton@gsu.edu). The authors thank the two anonymous *JM* reviewers for their constructive feedback and James Gibbs and Charlie Jordan for their comments on previous versions of the article.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

reflects the cultural experience of an ethnic minority group (an ethnic-oriented product) gains significant penetration among consumers outside the referent ethnic group.[1] Many marketers, though not all (cf. Clunis 2004), want to develop authentically ethnic-oriented products that reflect their referent culture well (i.e., that do not "sell out") but that also appeal to an audience that is sizable enough to ensure profitability (Andruss 2004; Guess 2004). However, the product attributes and other marketing decisions that give products their ethnic character may signal to mainstream and other-ethnic consumers that the products are not intended for them (Grier and Brumbaugh 1999). Alternatively, the product itself may be viewed favorably, but attitudes toward its consumption may be unfavorable. The purchase and consumption contexts surrounding an ethnic-oriented product likely affect whether it will cross over, and individual differences in preferences and beliefs about the referent ethnic group influence consumers' inclinations toward or away from ethnic-oriented products.

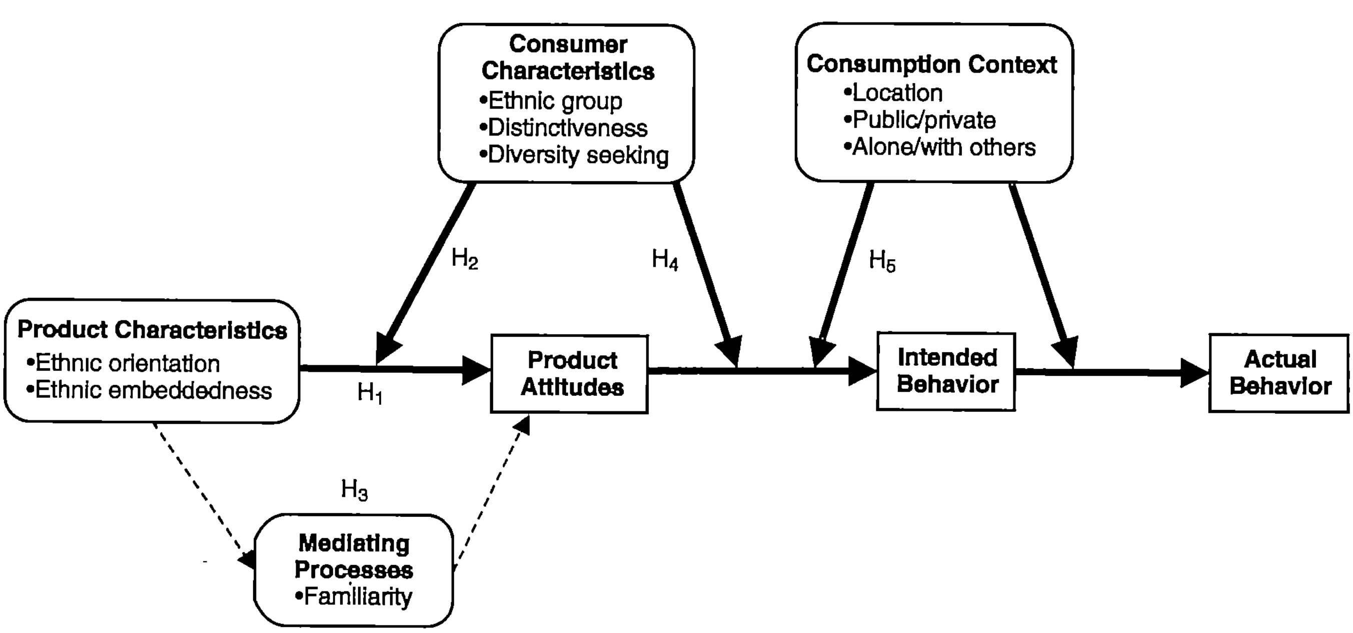
Our conceptual framework (see Figure 1) proposes that consumer characteristics and consumption context are important moderating effects on how ethnic-oriented product characteristics (including the impact of other marketing-mix decisions on inferred product characteristics) will be interpreted and embraced (or rejected) by mainstream or other-ethnic consumers. Important mediating processes influence how these product characteristics will be used to form attitudes, which influence subsequent behavioral intent and, ultimately, actual behavior.

This framework is based on attitude–intention–behavior models, which are frequently used to understand consumption intentions and behaviors across a variety of domains (Kozup, Creyer, and Burton 2003; Martin and Stewart 2001; Sheppard, Hartwick, and Warshaw 1988). Four additional key constructs, which we discuss in this section, are relevant for understanding crossover behaviors within this framework: characteristics that convey a product's ethnic orientation, consumer characteristics that influence how the product cues are interpreted, mediating processes that consumers use to form attitudes and behavioral intentions from product cues, and features of the context in which the product will be purchased and consumed.

Product Characteristics: Ethnic Congruence and Embeddedness

Some target market–specific product attributes preclude outright the product's intended use by non–target market members (e.g., feminine hygiene products' use by men).

FIGURE 1
Framework for Studying Ethnic-Oriented Product Crossover



Notes: Boldface indicates general constructs inherent in the generic framework. Bullet points indicate operationalizations we evaluated in Studies 1, 2, and 3 We note the specific hypotheses investigated as H_1–H_5.

---

[1] Several definitions are in order. We use the term "ethnicity," though in some cases the term "race" may be more accurate (our participants self-identified on the basis of racial categories of black or white). We are aware that the two terms have distinct meanings (i.e., ethnic group membership includes race but is not limited to this factor). Nonetheless, because some of our stimuli operationalizations are based on ethnic categories (African American), we use the term "ethnicity." In addition, this is conceptually consistent with prior research. The term "mainstream" refers to the numeric and social majority within a society, and the term "other-ethnic" refers to an ethnic minority group other than the referent ethnic group for the product in question. Throughout our studies, which take place in the United States, the referent ethnic minority group is black, the mainstream group is white, and other-ethnic groups include Hispanic and Asian American.

Similarly, some ethnic-oriented products function well only for members of the referent ethnic group (e.g., hair care products for blacks, Mandarin-language media for Chinese Americans). Consequently, crossover is not feasible for such products.

However, most ethnic-oriented products are not so constrained, and limits to crossover result from consumers' beliefs that the products are not intended or appropriate for them. Prior research shows that consumers respond less favorably to spokespeople of another ethnic group than to spokespeople of their own ethnic groups (e.g., Green 1999; Whittler 1991; Whittler and DiMeo 1991; Williams, Qualls, and Grier 1995) and suggests that people prefer products associated with their own ethnic group over products associated with other ethnic groups (cf. Whittler 1989).

Cues such as the ethnicity of a spokesperson or of the characters pictured on a package convey the ethnic orientation of a product. However, ethnic orientation is about not only the ethnic group associated with the product but also the degree to which the product is associated with that group, or its "ethnic embeddedness" (Williams 1995). Advertisements that combine ethnic characters with ethnic-consistent cues, such as depictions of ancestral heritage; culture-specific assimilation experiences; and other cultural-bound traditions, practices, and symbols, are highly ethnically embedded (Williams 1995). Such advertisements tap into the ethnic experience of ethnic-congruent consumers and engender favorable responses among target market members (Appiah 2001b; Brumbaugh 2002). In contrast, advertisements that are low in ethnic embeddedness include ethnic models combined with elements (e.g., background, language, themes) of mainstream American culture.

Products themselves may also contain ethnic-specific cues that influence the degree to which they are perceived as ethnically embedded. In addition, other marketing strategies, such as distribution in ethnic-frequented channels and promotion in ethnic-targeted media, may further signal to mainstream and other-ethnic consumers that an ethnic-oriented product is not intended for them and may reduce its potential for crossover. For example, tortillas, a product traditionally associated with Mexican ethnic culture, that are produced by machine, packaged with English-language labels by a well-known U.S. company, and distributed through national grocery store chains are far less ethnically embedded than tortillas that are rolled by hand by a small Mexican company, packaged with Spanish-language labels, and sold only in bodegas in the southwestern United States. Although the product category is associated with a specific ethnic group, other cues about its production, brand, country of origin, and distribution enhance (as in the latter example) or lessen (as in the former) its ethnic embeddedness.

Consumer Characteristics: Distinctiveness and Diversity Seeking

Characteristics of consumers influence how they interpret and respond to ethnic cues. Ethnic distinctiveness is one source of difference. Research on distinctiveness demonstrates that members of distinctive ethnic groups (groups that are a numeric minority and for whom ethnic group membership is salient and self-defining) respond more favorably to targeted appeals than do members of ethnically nondistinctive groups (groups that are a numeric majority and for whom ethnic group membership is neither salient nor self-defining) because of heightened levels of felt similarity that arise from ethnic congruence between product and consumer (Aaker, Brumbaugh, and Grier 2000; Grier and Deshpandé 2001; McGuire et al. 1978). For members of ethnically nondistinctive groups, ethnicity simply does not provide a basis for or information about attitudes toward targeted advertisements or products (Aaker, Brumbaugh, and Grier 2000).

Another characteristic relevant for ethnic crossover is consumers' propensity to seek out cultural or ethnic-related diversity in people, products, and services, a characteristic we call "diversity seeking." Related to but distinct from the brand-focused variety seeking ("the tendency of individuals to seek diversity in their choices of services or goods" [Kahn 1995, p. 139]), diversity seeking captures consumers' predilection for consumption experiences that are related to cultures other than their own, including those defined by ethnicity. Consumers who are not members of a particular ethnic group may be attracted to ethnic-oriented products because they seek variety outside the usual mainstream fare (Kahn 1993), because they have an interest in the specific culture (Appiah 2004), or because they are interested in different ethnic cultures in general (Halter 2000).

For example, Hallmark markets a collection of greeting cards called "Common Threads," whose messages reflect a variety of world cultures, emphasizing global community and diverse cultural expression (Halter 2000). Hallmark research found that "Common Threads" buyers were overwhelmingly white consumers who also tended to purchase multicultural or ethnic-oriented clothing, jewelry, and decorative home items. This segment is likely similar to Ray and Anderson's (2000) "Cultural Creatives." This group of people consists of 26% of adult Americans who are concerned with self-actualization, spirituality, and self-expression and who like things that are foreign and exotic.

A small but growing body of academic research has begun to explore inclinations toward cultural diversity. Miville and colleagues (Fuertes et al. 2000; Miville et al. 1999) developed a scale that measures people's tolerance of people who are different from themselves. Douthitt, Eby, and Simon (1999) developed a scale that assesses people's receptivity to dissimilar others. However, no academic research has empirically examined a diversity-seeking trait, whereby some people are highly interested in and proactively seek out people, places, products, and experiences that are culturally different from their own.

Consumption Context

Prior research shows that the contexts in which products are purchased and consumed have a profound impact on consumption beliefs and behaviors. For example, contextual factors dictate whether a product will be consumed at all (Cote, McCullough, and Reilly 1985) and, if so, what product attributes will be weighted more heavily in choice (Quester and Smart 1998). Social beliefs and norms influence consumption more for products that are consumed in public than for those that are consumed in private (Bearden and Etzel 1982), and perceived visibility of a product increases these social influences (Fisher and Price 1992).

The ethnic orientation of the context in which a product will be purchased or used has also been shown to influence its consumption. People are more likely to consume ethnic-consistent products when the consumption context is ethnically relevant than when it is associated with the mainstream or another ethnic group (Stayman and Deshpandé 1989). The presence of others of the same ethnicity increases the likelihood of ethnicity-consistent product consumption (Xu, Lotz, and Almeida 2004). Situations that heighten a person's ethnic-group salience lead to greater ethnicity-consistent consumption than situations that do not (Forehand and Deshpandé 2001).

Thus, it is highly likely that the specific context in which an ethnic-oriented product is used will influence the likelihood of crossover. Public consumption may increase the social risks of ethnic-oriented products for mainstream or other-ethnic consumers (Bearden and Etzel 1982; Fisher and Price 1992). When given a choice, mainstream or other-ethnic consumers may be more likely to use ethnic-oriented products in private consumption contexts (e.g., alone at home) than in public consumption contexts (e.g., with others at a public retail space). Similarly, they may be more likely to use ethnic-oriented private goods than ethnic-oriented public goods (Ratner and Kahn 2002; Richins 1994).

The nature of the other people with whom a person might consume ethnic-oriented products may further facilitate (e.g., if the others are of the referent ethnic group or are known for their tolerance of it) or inhibit (e.g., if others are not of the referent ethnic group or are known for their intolerance of it) the consumption of such products (Wakefield and Inman 2003; Xu, Lotz, and Almeida 2004). For example, the idea of an African aerobics class may appeal to some nonblack fitness enthusiasts. However, the thought of attending the class as a minority member or engaging in unfamiliar movements may undermine otherwise favorable product-based attitudes (toward the class itself), causing behavior-based attitudes and intentions (toward enrolling in and attending the class) to be far less favorable. Attending the class with a friend who is a member of the referent ethnic group may mitigate this effect.

Mediating Processes: Familiarity

Familiarity with the target product is one process that may mediate the impact of ad and product characteristics on attitudes and thus enhance the crossover of ethnic-oriented products. For example, familiarity gained from sampling Asian cuisine in restaurants is credited for increasing sales of Asian ingredients for use at home (*Retail World* 2000). Acceptance of ethnic desserts, such as Indian *kulfi* and Spanish *churros*, is greater when consumers are already familiar with similar offerings, such as ice cream and donuts (Strong 2002).

Thus, consumers' enjoyment of ethnic-oriented products likely depends on the ethnic-specific knowledge they have for such products. For example, whereas little ethnic-specific knowledge is required to understand mainstream-themed products, significant ethnic-specific knowledge is required to decode, interpret, and understand products whose meanings are culture bound with a specific ethnic group (Brumbaugh 2002; Grier and Brumbaugh 1999). If relevant ethnic knowledge is more difficult to access, as when a consumer is not a member of the referent ethnic group, the resulting lack of familiarity with the product may lower appeal for the product.

Studies Exploring the Framework

In summary, our studies focus on four key constructs that influence whether and to what degree an ethnic-oriented product will cross over: product characteristics, including marketing-mix factors that consumers use to make inferences about the product; consumer characteristics, such as diversity seeking; mediating processes, such as consumer familiarity with product cues; and situational factors that make a consumption context more or less conducive for consuming ethnic-oriented products. We explore various aspects of our framework in three studies that examine the crossover of ethnic-oriented popular films into the larger mainstream market. Film represents a vivid, influential, and economically significant social institution (Wolf 1999) in which ethnic-oriented products must cross over for economic viability (Acapulco Black Film Festival 1997; Alexander 2001; Bannon 2000). In all three quasi experiments (we measure participant ethnicity but do not manipulate it), we provide study participants with descriptions of movies that vary in their ethnic congruence with the participant and in their level of ethnic embeddedness. In Study 1, we explore how diversity seeking and ethnic distinctiveness (consumer characteristics) moderate participants' attitudinal responses to ethnic-congruent and -incongruent movies (product characteristics). In Study 2, we investigate the role of familiarity (a mediating process) on the influence of the product's ethnic congruence on attitudinal responses. In Study 3, we examine behavioral intentions among high- and low-diversity-seeking mainstream white participants (consumer characteristics) for different levels and types of ethnic embeddedness and situations (consumption context) associated with ethnic-oriented movies. We find support for our framework across all three studies, which we detail in the sections that follow.

Study 1

In Study 1, we explore the moderating impact of two consumer characteristics, ethnic distinctiveness and diversity seeking, on the main effect of congruence between consumer ethnicity and a product's ethnic orientation on its appeal. Prior research shows that people who perceive themselves as similar to ad spokespeople hold more favorable attitudes toward the advertisement than people who perceive themselves as dissimilar (Whittler 1991). However, research on target marketing shows that this holds only for ethnically distinctive (versus nondistinctive) people (Aaker, Brumbaugh, and Grier 2000). We expect to find a similar outcome for ethnic-oriented products (cf. Whittler 1989). Specifically, we hypothesize the following:

H_1: Ethnic congruence between a product and a consumer favorably influences attitudes toward the product only for members of a distinctive ethnic group.

Furthermore, we propose that people who are high in diversity seeking should be more interested in ethnic-incongruent products and find them more appealing than people who are low in this characteristic. We predict that diversity seeking will have a compensating effect on product preferences, enhancing attitudes toward a product only when participant ethnicity and product ethnic orientation are different.

H_2: People who are high in diversity seeking have more favorable attitudes toward ethnically incongruent products than people who are low in diversity seeking.

Method

Stimuli development and pretesting. We created three movie descriptions that represented three movie genres as replicates for our ad stimuli (for the stimuli for all studies, see Appendix A). We pretested each (n = 47) for consistency with the intended genre compared with the other two movies. The focal manipulation was movie ethnic orientation as conveyed through movie cues in each movie description. In addition, we pretested each manipulation for consistency with the intended ethnic group compared with other ethnic groups on seven-point scales. We manipulated movie ethnic orientation in *My Turn Again* (action genre) through the music of the main protagonist; "rock" was more associated with white ethnicity (M = 5.54) than "R&B" (M = 3.39; F(1, 46) = 31.56, $p < .01$), and R&B was more associated with black ethnicity (M = 4.65) than rock (M = 2.17; F(1, 46) = 39.55, $p < .01$). We manipulated movie ethnic orientation in *Thin Blood* (drama genre) through the name of the main protagonist; "Nathaniel Wellington" was more associated with white ethnicity (M = 5.17) than "Tyrone Washington" (M = 2.58; F(1, 46) = 56.18, $p < .01$), and Tyrone Washington was more associated with black ethnicity (M = 5.71) than Nathaniel Wellington (M = 3.13; F(1, 46) = 65.57, $p < .01$). Finally, we manipulated movie ethnic orientation in *The One* (romance genre) through the city and ethnicity of the main protagonists; "women in Seattle" were more associated with white ethnicity (M = 4.92) than "African American women in Detroit" (M = 1.17; F(1, 46) = 189.17, $p < .01$), and African American women in Detroit were more associated with black ethnicity (M = 6.83) than women in Seattle (M = 3.38; F(1, 46) = 153.79, $p < .01$).[2]

Design, procedure, and participants. Half of the participants saw the black version of *My Turn Again*, the white version of *Thin Blood*, and the black version of *The One*; the other half saw the other versions of the movies. Thus, the basic design of this study was a 2 (movie ethnic orientation: black or white) × 2 (participant ethnicity: black or white) × 2 (diversity seeking: high or low) mixed-factorial design in which movie ethnic orientation is a within-subjects factor and participant ethnicity and diversity seeking are between-subjects factors. Preference for genre and movie replicate factors are included as covariates.

Participants were 67 black and 62 white Americans who were recruited from public areas in two ethnically diverse cities on the West Coast. Participants ranged in age from 18 to 60 years; 59% were female. They were given a study packet that contained all stimuli and measures and were asked to read each movie description and to complete items that assessed their attitude toward the movie. After completing these measures for all three movies, participants completed items on their preference for different movie genres, their preference for movies about people different from themselves, and their demographic characteristics, including ethnicity.

Measures. We selected six movie attitude items on the basis of a review of prior industry marketing research that assessed the appeal of major movies among ethnically diverse Americans ("I would be interested in seeing this movie," "This movie would be entertaining," "This is my kind of movie," "This movie would be interesting," "The storyline attracts me to the movie," and "The characters attract me to the movie"). We assessed the six items on seven-point Likert scales and averaged the responses to yield a single dependent measure of attitude toward the movie ($\alpha = .95$). We used the Likert scale item "I like movies about people different from me" to create a diversity-seeking dummy variable containing zero or one to reflect the median split on this item (for additional details on this item, see Appendix B).[3]

Results

Unless otherwise indicated, degrees of freedom are (1, 386). The genre covariate was significant and positive. The movie covariate was significant such that *The One* was preferred over *Thin Blood*, which was preferred over *My Turn Again*. The model was significant (F(10, 386) = 8.99, $p < .01$). Both movie ethnic orientation (F = 6.90, $p < .01$) and participant ethnicity (F = 8.13, $p < .01$) were significant but were subsumed by a significant interaction between the two (F(2, 386) = 13.64, $p < .01$), as we expected. The main effect of diversity seeking was significant (F = 4.32, $p < .04$), and the anticipated three-way interaction among movie ethnicity, participant ethnicity, and diversity seeking was significant (F(2, 386) = 5.76, $p < .02$). We tested the hypotheses through a priori contrasts evaluated with one-tailed tests at $p < .05$.

H_1 proposed that movies for which movie ethnic orientation and participant ethnicity were congruent would be more appealing than movies for which ethnicity was incongruent, but only for members of distinctive ethnic groups.

---

[2] We recognize that this manipulation is not strictly parallel in that we do not specify "white" in the white manipulation (i.e., women in Detroit). We decided not to emphasize ethnicity for this condition because in the United States, white is the cultural and ethnic default in contrast to ethnic minorities (Brumbaugh 2002) and is rarely (if ever) used to describe the mainstream (Billig 1997).

[3] The results were similar whether the genre and diversity preference measures were included as continuous variables or as the median split dummies discussed here. We used the dummy variables to facilitate the interpretation and description of the results.

The results show that distinctive black participants preferred black-oriented movies (M = 4.44) over white-oriented movies (M = 3.55; F = 20.68, $p < .01$) and that nondistinctive white participants had no preference for white-oriented movies (M = 3.59) over black-oriented movies (M = 3.45, $p > .20$), as we hypothesized.

H_2 proposed that people who are high in diversity seeking would like ethnically incongruent movies more than people who are low in this characteristic. This held true for both white and black participants. Specifically, white participants liked black-oriented movies more when they were high in diversity seeking (M = 3.77) than when they were low in diversity seeking (M = 3.05; F = 4.56, $p < .03$); diversity seeking had no impact for ethnic-congruent white-oriented movies (preference high: M = 3.75; preference low: M = 3.48, $p > .20$). Black participants liked white-oriented movies more when they were high in diversity seeking (M = 3.91) than when they were low in diversity seeking (M = 3.05; F = 5.64, $p < .02$); diversity seeking had no impact for ethnic-congruent black-oriented movies (preference high: M = 4.46; preference low: M = 4.40, $p > .20$).

Discussion

We found support for both H_1 and H_2. As we anticipated, a match in ethnicity between potential viewers and movie actors enhanced attitude toward the movie only for distinctive black participants. Diversity seeking worked as hypothesized among all participants, enhancing attitude toward the movie only when movie and participant ethnicity were incongruent. The results are consistent with advertising research that shows that white consumers have no preference for white versus black sources featured in advertisements. However, these results are inconsistent with industry research that shows that white consumers tend to prefer mainstream or white-oriented movies over black-oriented movies (Easton 1991). To explore further what might be driving this discrepancy between our attitude results and observed phenomena, in Study 2, we evaluate the role of perceived familiarity as a mediating evaluation process between interpretations of ethnic-specific product cues and a product's appeal.

Study 2

People differ in their ability to make sense of ethnic-specific information when forming attitudes about ethnic-oriented products. Some may interpret a product's ethnic-specific cues in ways that make the product seem familiar to them, and others may not. Similarly, viewers' enjoyment of ethnic-oriented movies likely depends on the ethnic-specific knowledge they bring to bear on interpreting such movies. For members of the referent ethnic group of an ethnic-oriented movie, ethnic-relevant information is readily available and accessible. The more accessible and easily retrieved this information, the more familiar the viewer should be with the story depicted in the movie (Tversky and Kahnemann 1974) and the more the viewer should like the movie (Hoyer and Brown 1990; Laroche, Kim, and Zhou 1996; Loken and Ward 1990). However, such ethnic-specific knowledge may be either unavailable or more difficult to access for mainstream or other-ethnic people, and consequently, the content of ethnic-oriented movies may be more difficult to understand and less familiar. This lack of familiarity with the meaning of the movie's message or story line will likely result in lower appeal for such a movie. Therefore, we investigate the following hypothesis regarding the mediating influence of familiarity with an ethnic-oriented product on attitudes toward that product:

> H_3: (a) Ethnic congruence between a product and a consumer makes the product seem more familiar to the consumer, and (b) familiarity with a product enhances attitudes toward that product

Method

Stimuli development and pretesting. We created a single movie description for a comedy titled *Coming Home*. Movie ethnic orientation was conveyed through the ethnic group of the main protagonists (see Appendix A). Pretesting (n = 33) identical to that for the first study showed that the cue "Irish American" was more associated with white ethnicity (M = 5.36) than the cue "African American" (M = 3.71; F(1, 32) = 8.54, $p < .01$) and that African American was more associated with black ethnicity (M = 4.13) than Irish American (M = 1.97; F(1, 32) = 17.81, $p < .01$).

Design, procedure, and participants. This study had a 2 (movie ethnicity: black or white) × 2 (participant ethnicity: black or white) fully crossed between-subjects design that included genre preference as a covariate. Participants were 67 black and 49 white American adults who were recruited from the participant pools at two universities, one on the West Coast and one in the Southeast. Participants ranged in age from 18 to 45 years; 62% were female. Participants were randomly assigned to either the black- or the white-oriented movie condition. They first read a description of a movie and then responded to questions about their attitude toward the movie, familiarity with the movie, genre preference, and demographics.

Measures. After reading the movie description, participants completed the same six movie attitude items as those in Study 1 ($\alpha = .95$). In addition, they answered three seven-point Likert items that assessed their familiarity with the movie theme ("The movie would involve situations familiar to me," "The movie would be relevant to issues in my life," and "The movie would address topics I have experienced"); we averaged the responses to yield a single familiarity measure ($\alpha = .86$). A single item that was identical to the one in Study 1 assessed participants' preferences for the comedy movie genre.

Results

Unless otherwise indicated, degrees of freedom are (1, 115). The analysis of variance (ANOVA) model was significant for attitude toward the movie (F(4, 115) = 2.74, $p < .03$) and approached significance for familiarity (F(4, 115) = 1.92, $p < .11$). Genre preference was significant and positive in both models. The interaction between movie ethnic orientation and participant ethnicity was significant for familiarity (F = 6.10, $p < .02$). We tested the hypotheses at $p < .05$ through a priori contrasts evaluated with one-tailed tests.

Black participants preferred the black-oriented movie (M = 3.86) over the white-oriented movie (M = 3.35; F = 2.49, $p < .06$), but white participants did not prefer the white-oriented movie (M = 3.50) over the black-oriented movie (M = 3.55, $p > .20$). Thus, initial results of Study 2 support H_1 and are consistent with those of Study 1. H_{3a} posited that ethnic congruence between the viewer and movie ethnic orientation would enhance familiarity of the movie for the viewer. As we expected, black participants felt more familiar with the black-oriented movie (M = 3.83) than with the white-oriented movie (M = 2.96; F = 4.74, $p < .02$), and white participants felt marginally more familiar with the white-oriented movie (M = 3.77) than with the black-oriented movie (M = 3.12; F = 1.93, $p < .08$). Furthermore, H_{3b} is supported by a univariate regression, which shows that familiarity had a favorable impact on attitude toward the movie ($\beta = .47, p < .01$). As we anticipated, this effect was moderated neither by participant nor movie ethnicity, nor by their interaction (*ps* > .20).

Because ethnic congruence had a direct impact on attitudes for black participants per our reassessment of H_1, it is possible to evaluate whether familiarity mediates the impact of ethnic congruence on attitude toward the movie for black participants. To do so, we followed the three-step procedure that Baron and Kenny (1986) outline. The results for H_1 and H_{3a} show that ethnic congruence favorably influenced both attitude toward the movie and familiarity (the first step of their mediation test). Consistent with H_{3b}, familiarity positively influenced attitude toward the movie (the second step). If mediation occurs, the previously significant effect of ethnic congruence on movie attitude will be weakened when familiarity is included in the original ANOVA model (the third step). Indeed, black participants' preferences for the black-oriented movie over the white-oriented movie became nonsignificant when familiarity was included in the original model (F = .17, $p = .68$). Therefore, familiarity is a mediator of the impact of ethnic congruence on attitude toward the movie for black participants.

Discussion

Again, the results for black participants are consistent with expectations and previous research. Black participants felt more familiar with black-oriented movie cues and held more favorable attitudes toward the black-oriented movie than the white-oriented movie. Mediation analysis shows that familiarity mediated the impact of an ethnic match on attitude toward the movie among black participants. The results of the movie manipulation for white participants were unremarkable; familiarity was marginally affected by ethnic congruence.[4] A movie's ethnic orientation, when manipulated by actor ethnicity, ethnicity associated with the city in which the movie takes place, and ethnicity associated with the music genre related to the plot, has little impact on white participants' attitudes toward the movies to which they were exposed in the previous studies.

This result suggests that white consumers' attitudes toward ethnic-oriented movies would equally favorably dispose them toward seeing white- and black-oriented movies. Again, movie industry data contradict this interpretation. White consumers simply do not view black-oriented movies with the frequency that our attitude measures would imply (Easton 1991; Masters 1996). We offer two explanations for this observation: First, explicit attitude measures may be inappropriate because of social desirability and other biases. People may be unwilling to share their true feelings about other-ethnic movies (Devine 1989) or may adjust (consciously or subconsciously) their initial reactions toward more socially acceptable and less-biased responses (Martin, Seta, and Crelia 1990; Monteith 1993). Second, consistency between a person's overall attitude toward a movie and the actual behavior of seeing that movie at a particular theater with specific others may be particularly low (for a review, see Kraus 1995). Temporal distance between measuring attitude toward a movie and actual attendance is likely to weaken the relationship between attitude and behavior (Davidson and Jaccard 1979). The movie itself as attitude object may be unrelated to the behavior of attending the movie, attenuating the relationship between the two (Ajzen and Fishbein 1977).

To address these issues, we present a series of behavioral intention items in our final study to evaluate how our experimental factors influence intended behaviors in addition to attitude toward the movie. We also evaluate hypotheses that propose that some marketing-mix factors may enhance or detract from an ethnic-oriented movie's attraction among white consumers, depending on their level of diversity seeking.

Study 3

Most research shows that movies with a high level of ethnic embeddedness fare poorly with mainstream audiences, attract smaller audiences (Carter 2001), and make lower revenues (Alexander 2001) than movies with mainstream themes. Movies that are high in ethnic embeddedness are frequently produced by smaller, independent filmmakers with fewer resources to devote to hiring big-name stars, gaining access to distribution channels, and promoting their films (Acapulco Black Film Festival 1997; Alexander 2000; Johnson 2003). In addition to the overall ethnic-oriented theme of such movies, these other marketing-mix decisions likely reinforce or lessen the product's perceived ethnic embeddedness. Unknown actors do not have the draw that famous actors do (De Vany and Walls 1997), reducing possibilities for crossover. Smaller promotional budgets used selectively in ethnic-consistent media likely either fail to reach potential mainstream viewers or, if these viewers are exposed to promotions for the film, imply that the movie is not intended for them. Limited distribution in ethnic-oriented venues may limit consumption among mainstream consumers who are unaccustomed to being a numeric minority.

---

[4] Puzzled by the lack of effects on white viewers, who are crucial for crossover, we conducted another study that built on the first two by manipulating the presence or absence of a targeting cue ("These movie concepts have been designed to appeal to people like you as the target market") included with a movie description with the intent of enhancing felt targetedness, which has been found to influence attitudes among whites in previous research (Aaker, Brumbaugh, and Grier 2000). Again, we manipulated movie ethnic orientation as black or white through the R&B or rock music cue, and participants included black and white adults and students from university participant pools. As we expected, black participants had more favorable attitudes for ethnic-incongruent movies when the descriptions were accompanied by the targeting cue than when they were not. However, the targeting cue had no impact on targetedness or attitudes for whites' responses to the black-oriented movie, as was the case with our previous studies.

Therefore, in this final study, we evaluate two hypotheses about how marketing decisions influence mainstream consumers' behavioral intentions toward ethnic-oriented movies. Product factors that enhance a movie's ethnic embeddedness (e.g., actor ethnicity, actor fame, ethnic-oriented plot) may contribute to perceptions that a movie represents the referent culture well. People who are interested in authentic representations of other cultures (e.g., those who are high in diversity seeking) may be more likely to see films that are high in ethnic embeddedness than those who are not. Therefore, we hypothesize the following:

> H_4: Factors that contribute to an ethnic-oriented product's ethnic embeddedness enhance behavioral intent for that product among mainstream consumers, particularly those who are high in diversity seeking.

However, the same marketing-mix factors that enhance the ethnic embeddedness of a movie may also limit whether mainstream viewers are willing to see it. Specifically, lower marketing budgets, especially if used in targeted media, and limited distribution in ethnic-oriented venues not frequented by mainstream customers may reduce the likelihood that mainstream consumers will see the film, particularly if they were not otherwise favorably disposed toward the ethnic orientation of the movie (e.g., those who are low in diversity seeking). Therefore, we hypothesize the following:

> H_5: Factors that contribute to an ethnic-oriented consumption context reduce behavioral intent for ethnic-oriented products among mainstream consumers, particularly those who are low in diversity seeking.

Method

Stimuli development and pretesting. We created two detailed movie descriptions, *Killer Mind* and *The Story*, which included the plot, a review (same across conditions), and additional manipulated information on promotion and distribution of the film (see Appendix A). In the first set of the movie descriptions, we include names of black actors and systematically manipulate four two-level factors of interest: marketing budget (large or small), location where the movie was being shown (associated with black or mainstream audiences), ethnic embeddedness of the plot (high: associated with black culture; low: associated with mainstream culture), and actor fame (famous or unknown). We pretested (n = 57) the actors' names to ensure that they were either recognized as or inferred to be black actors and to ensure that they were either famous or unknown. Indeed, the two famous actors used in the study were recognized as such (M = 6.27) compared with the two fictitious actors used (M = 3.14; F(1, 227) = 360.24, $p < .01$). In a recognition task, 90.4% of participants recognized the names of the famous actors, and 95.6% did not recognize the names of the unknown actors ($\chi^2(1) = 169.00$, $p < .01$). These four black actors were also identified as black (M = 6.09) compared with the white (M = 1.82) and Hispanic (M = 1.52) actors we included as foils in the pretest (F(2, 683) = 852.95, $p < .01$).

Marketing budget was either $27 million (just over 2003's $25 million average; Friedman et al. 2003) in the large condition or $1.3 million (approximately 5% of the large budget; Alexander 2001) in the small condition. To make this manipulation more salient and meaningful, the text of the budget manipulation also includes the implications of a large or small marketing budget on advertising reach. The location of the movie was stated as either where black viewers would be likely to see it (black condition) or where a majority of viewers would be likely to see it (mainstream condition). To effect the plot manipulation, we included a subplot in the movie description that stated either that black ethnicity was of focal interest to the story line (high ethnic embeddedness) or that another facet of the subplot was of focal interest to the story line (low ethnic embeddedness).

In the second set of the movie descriptions, we created one mainstream control version for each movie in which there was no mention of actor name, race, or fame. The movie budget was large, the location was mainstream, and the ethnic embeddedness of the plot was low.

Design, procedure, participants, and measures. In this design, the movie descriptions comprising the four manipulated factors are nested under the black condition of a two-level actor ethnicity factor, and the mainstream control descriptions are nested under the mainstream condition of this factor. We used the SAS design procedure to generate an incomplete factorial design with the four between-subjects factors plus the movie replicate for the two movies that enabled us to assess all main effects and two-way interactions among design factors.[5] This design yielded 16 different conditions nested under black actor. Therefore, the design of this study was an incomplete 2 (actor fame: famous or unknown) × 2 (budget: large or small) × 2 (location: black or mainstream) × 2 (ethnic embeddedness of plot: high or low) plus control ANOVA model.

Participants were 97 white American MBA students who were recruited from a participant pool at a southeastern university; they participated in this study as one of several studies in a group testing session. Participants ranged in age from 22 to 33 years; 37% were female. Of the 97 participants, 86 were randomly assigned to read one of the 16 black-oriented movie descriptions, and 11 were randomly assigned to read both of the mainstream movie descriptions. All participants first read a description of the movie to which they had been assigned and then responded to the same questions about their attitudes toward the movie (α = .96) used in Studies 1 and 2. Then, they completed eight additional seven-point items that assessed the likelihood that they would see the movie in a theater or rent the video or DVD and whether they would do so alone, with a friend, with their spouse or partner, or with a work colleague. These questions comprise two additional fully crossed within-subjects factors: venue (see in theater or rent on video or DVD) and social context (alone, with friend, with spouse or partner, or with work colleague). Participants in the mainstream condition repeated this procedure for the second movie. Finally, participants provided measures of genre preference, diversity seeking, and demographic data, as in the previous studies.

---

[5] Power analysis suggests that we would require more than 500 participants to evaluate the fully crossed design. Given that we had no prior hypotheses regarding three- and four-way interactions, we elected to study a more parsimonious and practical model that enabled us to evaluate main effects and two-way interactions but that required only approximately 100 participants.

Results

We initially analyzed attitude toward the movie with the aforementioned ANOVA model that included all main effects and two-way interactions among design factors nested under actor ethnicity (black versus mainstream control) as well as the main effect of and interactions with diversity seeking. Although the model was significant (F(16, 96) = 3.00, $p < .01$), this significance was driven only by the covariates; no experimental design effects were significant ($ps > .17$). Again, movie attitudes for white consumers were unaffected by our experimental manipulations.

For behavioral intentions, we also crossed the ANOVA model with venue and social context factors. No four- or five-way interactions among design factors and diversity seeking, venue, and social context achieved significance, so we dropped them from the final analyses. We included attitude toward the movie as a covariate, and unless otherwise stated, degrees of freedom are (1, 775). We tested hypotheses at $p < .05$ through a priori contrasts evaluated with one-tailed tests. Relevant cell means appear in Table 1. This model was significant (F(117, 775) = 5.05, $p < .01$), and actor ethnicity had a main effect such that participants were more likely to see the mainstream control movies (M = 3.75) than the black-oriented movies (M = 3.22; F = 10.80, $p < .01$), consistent with industry data.

We hypothesized two effects on intended behaviors. First, in H_4, we predicted that characteristics of an ethnic-oriented movie that added to its ethnic embeddedness would increase the chances that white customers would see the movie compared with ethnic-oriented movies that were low in ethnic embeddedness, particularly for high diversity seekers. Indeed, plot and actor fame interacted (F = 5.71, $p < .02$) such that all participants were more likely to see a movie with a highly ethnically embedded plot and unknown actors (M = 3.67) than with a less ethnically embedded plot and/or famous actors (for the remaining three cells, M = 3.07; F = 3.25, $p < .04$). Furthermore, the significant three-way interaction among plot, budget, and diversity seeking was significant (F = 4.95, $p < .02$) such that low-diversity-seeking participants were less willing to see movies with highly ethnically embedded plots and large budgets (M = 2.48) than other movies (for the remaining three cells, M = 3.32; F = 5.70, $p < .01$), consistent with H_4. High-diversity-seeking participants were equally willing to see movies regardless of plot and budget conditions.

Second, in H_5, we predicted that other marketing factors (particularly distribution decisions) that increased the ethnic orientation of the consumption context would decrease the chances that white consumers would see an ethnic-oriented

TABLE 1
Study 3: Effect of Diversity Seeking and Movie Characteristics on White Participants' Likelihood of Seeing a Movie with Black Actors

	Diversity Seeking			
	Low		High	
Budget	**Small**	**Large**	**Small**	**Large**
Location				
Ethnic	2.48	3.13	3.32	3.31
Mainstream	3.91	3.10	3.29	3.16
Plot				
High embeddedness	3.60	3.10	3.61	3.27
Low embeddedness	2.48	3.13	3.02	3.23
	Actor Fame			
	Unknown	**Famous**	**Unknown**	**Famous**
Venue				
Rent	3.14	3.11	3.83	3.26
See	3.20	3.05	2.87	3.13
Social Context				
Friend	3.28	3.82	3.83	3.87
Partner or spouse	3.88	3.41	3.21	3.72
Self	2.28	2.18	2.71	1.93
Work colleague	3.25	2.91	3.64	3.26

movie, particularly among low-diversity-seeking consumers. The significant interaction between budget and venue (F = 3.99, $p < .05$) shows that participants were less likely to see a high-budget movie in a theater (M = 2.92; F = 9.51, $p < .01$) than to rent it (M = 3.47); there was no difference in whether they would rent (M = 3.26) or see in a theater (M = 3.22) small-budget movies ($p > .20$). The three-way interaction among location, budget, and diversity seeking was significant (F = 11.49, $p < .01$) such that low-diversity-seeking consumers were less likely to see a small-budget film at an ethnic location (M = 2.48) than at a mainstream location (M = 3.91; F = 15.75, $p < .01$) and compared with a high-budget film (for two cells, M = 3.11; F = 12.74, $p < .01$).

Actor fame interacted with venue and diversity seeking (three-way interaction: F = 3.23, $p < .07$) such that low-diversity-seeking participants were equally willing to rent or see movies regardless of whether the actors were famous (*ps* > .20). High-diversity-seeking participants were more willing to rent a movie with unknown black actors (M = 3.83) than they were to see a movie with black characters in the theater or rent a movie with famous black actors (for three cells, M = 3.10; F = 8.46, $p < .01$).

Finally, the person with whom participants would see a movie affected the types of movies they were willing to see. In the control condition, people were equally willing to see a movie with any other person, whether a friend (M = 4.27), a partner or spouse (M = 4.23), or a work colleague (M = 3.98), but they were less willing to see a movie alone (M = 2.52) than with these others (for three cells, M = 4.16; F = 44.30, $p < .01$). This was not true for high- and low-diversity-seeking participants seeing black-oriented movies with famous or unknown actors in different social contexts (three-way interaction: F(3, 775) = 2.55, $p < .05$). Low-diversity-seeking participants were more willing to see a movie with a friend if it had famous actors (M = 3.82) than if it had unknown actors (M = 3.28; F = 2.07, $p < .08$); this difference was not significant for high-diversity-seeking participants ($p > .20$). High-diversity-seeking participants were more willing to see a movie with a partner or spouse if it had famous actors (M = 3.72) than if it had unknown actors (M = 3.21; F = 7.08, $p < .01$); this difference was not significant for low-diversity-seeking participants ($p > .20$). High-diversity-seeking participants were more willing to see a movie alone if it had unknown actors (M = 2.81) than if it had famous actors (M = 1.93; F = 2.67, $p < .05$).

Discussion

The results provide support for H_4 and H_5 as well as insight into how high- and low-diversity-seeking white consumers interpret different cues and marketing tactics regarding ethnic-oriented products. As we predicted in H_4, all participants were more likely to see a black-oriented movie that had a highly ethnically embedded plot and a low promotional budget than to see other types of black-oriented movies. We interpret this result to suggest that something about the authenticity of the movie as conveyed through the plot and lack of commercial push was more appealing for white viewers than black-oriented films of the more mainstream comedic or "gangsta" genres. High-diversity-seeking participants also seemed to interpret the use of unknown actors positively, particularly for home rental, and were more likely to see black-oriented movies with unknown actors alone or with a friend than were low-diversity-seeking participants. In summary, white participants, especially those high in diversity seeking, were favorably disposed toward seeing black-oriented movies that contained several cues (versus only one or few) that reinforced their ethnic orientation.

Social norms seemed to have a different effect on low- and high-diversity-seeking participants. Notably, high-diversity-seeking participants were less likely to go with a partner or spouse to see a black-oriented movie that had unknown actors than one with famous actors, but low-diversity-seeking participants were less likely to go with a friend to see a black-oriented movie that had unknown actors than one with famous actors. Examination of these means in Table 1 suggests that the social risk associated with seeing a black-oriented film with unknown actors was higher with more casual acquaintances than with a partner or spouse for low-diversity-seeking participants. In contrast, high-diversity-seeking participants either were not bothered by such risk or had friends who were similarly favorably disposed toward black-oriented movies with unknown actors. Instead, high-diversity-seeking participants were less likely to see that movie with a partner or spouse who, perhaps, had little inclination toward such films.

General Discussion

The findings of our studies are consistent with prior research with regard to black participants. The results across the first two studies show that black participants preferred black movies over white movies, echoing work on advertising response. The diversity-seeking characteristic enhanced attitudes toward white movies among black participants, compensating for the absence of a match in ethnicity. Furthermore, we demonstrated that though familiarity positively enhanced attitudes for all participants, familiarity mediated the impact of ethnic congruence on attitudes only for black participants. These results enable us to understand more clearly why mainstream movies attract ethnic minority viewers: Familiarity with their dominant cultural themes enhances attitudes for potential ethnic minority viewers, and ethnic minority viewers who were high in diversity seeking were more favorably disposed toward other-ethnic mainstream movies.

The results among white participants also demonstrated that the diversity-seeking characteristic moderated whites' attitudes, such that white participants who were higher in diversity seeking had more positive attitudes toward black movies than did white participants who were lower in diversity seeking. Although the main effect of the ethnic orientation of the movie did not directly influence attitudes among white participants, it did affect behavioral intentions: The presence of black actors decreased the likelihood that white participants would see the movie compared with mainstream control conditions. The results show that marketing-mix factors interacted with the social context of product consumption to influence the behavioral intentions of whites toward black movies.

Implications for Practice

The crossover of ethnic-oriented products from niche segments into larger consumer segments may have positive social and economic consequences. Ethnic-oriented products such as movies serve as socialization vehicles, exposing people to a variety of images, cultures, and lifestyles not otherwise learned through personal experience (Gerbner et al. 1994; Tharp 2001). However, many ethnic-oriented products, particularly those reflecting U.S. ethnic minorities, are targeted solely toward the referent ethnic group because of concerns that other ethnic groups will be uninterested (Carter 2001; Johnson 2003; Masters 1996). Our results identify strategies that marketers of ethnic-oriented films and other ethnic-oriented products might use to generate crossover appeal along with social and economic benefits.

Our results show that ethnic congruence enhances familiarity and that familiarity enhances attitudes among all consumers. Changing product characteristics to enhance appeal among target viewers is a standard movie industry practice (Diamond 1989; Pampel, Fost, and O'Malley 1994) that may be used fruitfully to increase crossover appeal of ethnic-oriented products among mainstream consumers. However, it may be more difficult (or undesirable) to change the characteristics of highly ethnically embedded movies whose story lines are inextricably bound to ethnic experiences simply to elicit feelings of familiarity among white audiences. Alternatively, marketers might induce feelings of familiarity through promotional changes (versus product changes) that focus on aspects of the movie or product that either transcend ethnic boundaries or hold unique appeal among relevant subsegments. Our results indicate that more concerted attention toward the creation of strategic crossover with multiple targeted promotional messages placed in carefully selected media outlets may prove successful.

The findings also demonstrate that there is a subset of white viewers who may seek out cultural diversity in some of the products they choose. Marketers of ethnic-oriented films and other ethnic-oriented products may find success in inducing crossover in regions where high-diversity-seeking white consumers are located. Geodemographic research shows that there are distinct neighborhoods that are increasing rapidly in ethnic composition while retaining diversity across a range of ethnic groups, including whites. Places such as Orange County, Calif. (Yi and Herndon 2001), and Queens, N.Y. (Campos-Flores 2001), have favorable socioeconomic characteristics and critical masses of ethnic-group consumers. Whites who choose to live in these areas may well be those who seek out and value diversity. Such geographic concentration is ideal for distribution-based strategies to increase crossover potential of ethnic-oriented products such as films. For example, by placing highly embedded ethnic-oriented movies in such neighborhoods, marketers may efficiently and effectively target whites who are predisposed toward attending them.

The results also demonstrate that the relationship between the degree of the ethnic embeddedness of a product and whites' behavioral intentions toward that product is not straightforward. Marketers of ethnic-oriented products sometimes attempt to water down their offerings to make them more palatable for mainstream or other-ethnic audiences. However, our findings indicate that this strategy may be unsuccessful for two reasons: First, low-diversity-seeking white consumers may be uninterested in an ethnic-oriented product regardless of its degree of ethnic embeddedness simply because its ethnic orientation alone signals to these consumers that the product is not for them and thus reduces its appeal. Second, weakening the ethnic embeddedness of the product may eliminate the very characteristic that appealed to high-diversity-seeking white and referent ethnic consumers in the first place, namely, genuine representation of the referent ethnic group. This counterintuitive (and even ironic) nature of the degree of ethnic orientation was evident in white participants' behavioral intention ratings. Highly ethnically embedded films tend to have unknown actors, lower budgets, and sparser distribution, whereas less ethnically embedded films (mainstream ethnic films) tend to be supported with well-known actors, larger budgets, and more widespread distribution. As the typical ethnic fare, the mainstream ethnic films appeared to hold less appeal to whites in our sample, perhaps because such movies are perceived as no different from other mainstream movies and thus evoke no special interest. Ironically, these are most likely the ethnic movies that whites are made aware of through greater promotion and distribution efforts. Given that, in general, highly ethnically embedded films evoke greater appeal, especially among high-diversity-seeking whites, marketers may want to create different marketing strategies depending on how deeply the product-related cues are embedded in the referent ethnic culture and how these cues resonate with both referent ethnic consumers and low- and high-diversity-seeking mainstream consumers.

Our findings similarly demonstrate that the strength of association of a particular ethnic group with a consumption context may also facilitate (or hinder) the consumption of ethnic-oriented products. Thus, the results emphasize the importance of considering not only how widely a product is distributed but also where the product is distributed and the ethnic meanings associated with the distribution context. Although changing the ethnic composition at or beliefs about an ethnic-specific venue or retail location may be challenging, creative promotional strategies can be used to overcome contextual barriers to consumption. For example, special events at the consumption venue that appeal to mainstream or other-ethnic audiences, sales promotions that encourage consumption among mainstream or other-ethnic audiences, and communications that signal that mainstream and other-ethnic consumers are especially invited to partake of ethnic-oriented offerings might ameliorate distribution-specific causes of crossover failure. In general, the results emphasize the significance of the ethnic composition of consumption contexts, the role of a person's companion in facilitating or inhibiting other-ethnic product consumption, and the ethnic embeddedness of the product in attracting non–target market consumers.

Theoretical Contributions

Our framework identifies key constructs that are relevant to the crossover of ethnic-oriented products and examines the relationships among these factors that affect an ethnic-oriented product's ability to garner interest in and gain sales among mainstream and other-ethnic consumer segments. The framework generalizes beyond our application to white viewers' predilection for black-oriented movies to other product categories, other referent ethnic groups, and other nontarget customer segments. Because services possess a high degree of social embeddedness (i.e., services are delivered by people and frequently produced and consumed in the presence of other people), they may provide a particularly rich context in which to extend this framework. For example, the framework would help explain whether or not a white mother would take her biracial daughter with "frizzy" hair to an African American–oriented salon located in an African American neighborhood. The framework also provides a mechanism for understanding how the social, cultural, and historical beliefs about other referent ethnic groups (e.g., Asian, Hispanic) likely influence the crossover of products associated with these groups into mainstream and other-ethnic segments. Extending the framework further, it similarly provides an approach for understanding how to market a product that has become strongly associated with a particular consumer segment on the basis of characteristics other than ethnicity (e.g., age, gender, sexual orientation, or region) to a broader group of consumers.

Our results also contribute to the theoretical understanding of the processes by which non–target market consumers form inferences about products that are associated with specific consumer segments. One clear outcome of this research regarding such processes is the inadequacy of attitudes in predicting whites' behavioral intentions toward ethnic-oriented products. In general, it is accepted in consumer contexts that attitude toward a product is somewhat indicative of purchase intent (Kraus 1995). However, research on the attitude–behavior link shows that this link may be attenuated by the social characteristics of potential consumption contexts (Zalesny and Ford 1990). We believe that social context explains our findings similarly. Although whites' attitudes were seemingly unaffected by the ethnic orientation of a movie, it was not because ethnicity does not matter to these white consumers. Rather, our attitude measures were decontextualized and did not reflect or include the appreciable influence of ethnic-oriented context on movie consumption behavior. Moviegoing is fundamentally a social activity (Austin 1989), and our results show that the ethnic orientation of the social context in which a person anticipates seeing a given movie has a profound impact on the decision to see that movie. Further research exploring a range of dependent measures that tap into purchase intention and behaviors, perhaps implicitly to avoid social desirability biases, might show greater variability among and insights into white consumers and ethnic-oriented products.

In addition, our findings extend burgeoning research on ethnic embeddedness (Appiah 2004; Brumbaugh 2002; Williams 1995) that has examined the degree to which advertisements contain ethnic cues. Our studies examine not only the ethnic embeddedness of the product itself but also its consumption context. Our results are consistent with those of Appiah (2001a), who finds that some products associated through advertising with black culture may be desired by mainstream consumers more than products associated with mainstream culture. The associations related to culture alter the nature of the products themselves, giving them different meaning that causes them to resonate with certain, possibly unexpected, segments. Ethnic-oriented films that are high in ethnic embeddedness and that possess qualities that are authentic to the ethnic-group culture seem to be endowed with a similar ability to resonate with high-diversity-seeking consumers. Thus, our work extends the limited research on ethnic embeddedness in the advertising domain into products and consumption contexts.

Limitations and Directions for Further Research

As in all research, we made some concessions that create limitations in this research and opportunities for future study. Across all three studies, we assessed consumers' responses to descriptions of movies rather than to filmed movie trailers. Although certainly not as rich as multimedia clips or advertisements, such descriptions are common in newspapers and television guides and serve an advertising function in and of themselves. Research shows that people use such descriptions to make moviegoing decisions (Cooper-Martin 1991), and industry research before final production relies on similar stimuli (Austin 1989). Nonetheless, additional investigations with movie advertisements and trailers could further increase the generalizability of our results.

A potential problem with our operationalization of movie venue location in Study 3 is its implied distance from our white participants' homes. Specifically, our white participants may have inferred that they not only would be in the minority at the black movie venue but also would need to travel farther to reach that venue compared with the mainstream locations. Although follow-up analyses preclude this explanation,[6] we readily concede that this is a limitation of Study 3 and encourage future researchers to test these contextual effects on crossover potential.

Our findings reveal subtle but important interactions among marketing-mix factors that influence whites' willingness to see ethnic-oriented movies. However, because of the complexity of the design and the number of factors we explored in Study 3, we were unable to investigate all combinations of marketing and movie characteristics that might increase white consumers' likelihood of seeing ethnic-oriented movies in theaters. For example, simply increasing the marketing budget to make high-diversity-seeking consumers aware that a highly ethnically embedded movie was showing in their area might increase their attendance. This effect is captured by a five-way interaction that we were unable to evaluate. Additional research could explore these interactions and relationships.

---

[6] It has been noted that the black/mainstream location factor may be confounded with perceived distance such that participants are less likely to see a movie at a black location than a mainstream location because they perceive the former to be farther away than the latter. Additional analyses suggest that this is not the likely cause. We analyzed all of our behavioral items (whether participants would see the movie in a theater under different conditions), with our covariates and location as our sole factor. Multivariate ANOVA results show no difference in means across dependent measures between black and mainstream locations ($p > .20$). Univariate results show that only the difference of whether a person would see the movie with a friend at a black versus a mainstream location approached significance ($p < .06$). Therefore, we conclude that participants in this study are unlikely to associate the ethnic orientation of the location with distance.

Our results indicate that diversity seeking may be a fruitful segmentation variable, especially given the increase in ethnic-oriented products and the changing ethnic composition of the marketplace. In Studies 1 and 3, we used a context-specific, single-item measure that we assessed after presenting our movie stimuli. Both the use of a single item and the timing of its administration in our studies are weaknesses that could be remedied in further research. Although we provide support for this single item (see Appendix B) and believe that it is an adequate indicator of diversity-seeking consumption behavior in this moviegoing context (Drolet and Morrison 2001), additional studies might use existing related scales (Douthitt, Eby, and Simon 1999; Miville et al. 1999) to evaluate the extent to which this characteristic influences consumption of ethnic-oriented products. Alternatively, further research might build on our steps toward the development of a context-neutral measure of diversity seeking related to consumption activities to identify consumers who are predisposed toward products associated with other ethnic groups. These initial explorations indicate that this characteristic may have multiple dimensions that influence distinct consumption behaviors in different ways. Our single context-specific item scratches the surface of what may be a rich area for future exploration.

We conjectured in the discussion of our final study that consumers might use product cues to assess the authenticity of an ethnic-oriented movie. We suggested that highly ethnically embedded plots with lesser-known actors and lower marketing budgets could be interpreted as more representative of the ethnic cultures portrayed and that this authenticity makes such movies more attractive to high-diversity-seeking consumers. Authenticity is indeed an inference that marketers attempt to create in their target markets for ethnic products (e.g., Watts and Orbe 2002). Further research should address whether consumers make such assessments of authenticity and whether such assessments mediate (as our discussions imply they might) the impact of marketing tactics on behavioral intentions (cf. Grayson and Martinec 2004).

Finally, our research examined the traditional notion of crossover from a niche market into the mainstream market. However, a key to increasing the economic viability of ethnic-oriented products may be consumption by members of other ethnic minorities, who may identify with the referent ethnic minority's experiences and may be predisposed toward products associated with the referent ethnic minority group. Research that includes multiple ethnic groups will provide additional theoretical and practical insights into crossover among a variety of consumer segments.

Our framework provides a robust structure in which to study diverse consumer segments' consumption of an equally diverse portfolio of goods and services (Hirschman and Holbrook 1982). In applying this framework to movies, we have begun to understand how mainstream white consumers perceive ethnic-oriented products. Additional studies that apply this framework to the crossover of other product categories, among different consumer segments, with other social characteristics and individual difference factors, and across a variety of consumption contexts will help marketers understand, predict, and enable crossover consumption in a multicultural marketplace.

Appendix A
Movie Descriptions

Study 1

My Turn Again (action genre). In this action-packed film, Tony Larkins, a **rock/R&B** music singer declining in popularity, marries a shy country girl and helps her become a famous actress. Her fame not only eclipses his, but he gradually realizes that he was used and plots his revenge.

Thin Blood (drama genre). In a modern-day romance, **Nathaniel Wellington/Tyrone Washington**, a devoted son, must choose between caring for his cantankerous but well-meaning father and moving across the country to marry the woman he loves. While his mother wants him to stay near home, his sister, who fell out of her father's favor, argues that he should do what he wants. Nathaniel is torn, and his uncertainty, although comic at times, just might lead to tragedy.

The One (romance genre). This comedy finds two single professional **women in Seattle/African American women in Detroit** searching for *The One* in all the wrong places! Susan is a lawyer, and Jackie is a business school professor. Their time-consuming jobs mean they must make their search as "efficient" as possible. So they decide to place a "highly creative" singles ad in the local paper. Their creativity is surpassed only by that of the variety of men who respond to their ads.

Study 2

Coming Home. This light comedy finds three **Irish American/African American** brothers suddenly living under the same roof for the first time since childhood. Their different experiences have changed each of them in ways unimaginable to the others. Each day brings new discoveries and excitement for the brothers, as they each pursue careers, women, and a good time on the town!

Study 3

The Story. The main character in the film is a career-focused, up-and-coming newscaster at a local morning news show who believes she has everything going for her. When studio execs at a competing show take a possible interest in hiring her, her boss decides to get her out on the streets to start covering stories herself to see if she has what it takes to

TABLE A1
Factor Loadings for Diversity-Seeking Subscales

	Learning	Hiring	Eating	Living
Learning about other cultures is something I enjoy.	.8275	.0835	.2670	.3672
I view travel as an opportunity to learn about different cultures.	.7608	.0142	.2265	.0796
When choosing where to go on vacation, I favor places with people of different ethnic groups.	.5017	.2958	.3614	.2065
I think too much emphasis is placed on diversity in hiring decisions. (R)	.1040	.7768	.0890	.2220
If I were hiring, I would try to seek out candidates who are different from the current work force in terms of ethnicity, race, gender, and sexual orientation.	.0458	.6673	–.1134	.3442
I think it is important to create a workplace that represents the local population in terms of ethnicity, race, gender, and sexual orientation.	.0548	.5892	.1926	.1687
I enjoy trying different types of food from foreign countries.	.3211	.0475	.6945	.0384
I avoid going to restaurants with foods from novel places (e.g., Ethiopian or Peruvian). (R)	.1202	.1004	.6582	.2363
Other than Chinese or Mexican food, I don't really seek out "ethnic" foods. (R)	.3188	.0545	.6381	.1183
I would welcome a member of another cultural, ethnic, or racial group into my family.	.1409	.2137	.0746	.6879
It is very important to me to live in a neighborhood with people from different ethnic groups.	.2459	.4024	.2011	.5808
I would consider living in an area where there lived many people who were not from my ethnic group	.1784	.2578	.2125	.5783

Notes: R = item is reverse scored.

be a serious journalist for the station. She is none too happy, however, when she is assigned to the best cameraman in the crew—the two have never gotten along, and sparks fly the minute they start working together. As it turns out, it is the first time a **Black/this** newscaster and cameraman have been paired together, and the comedic **racial-oriented/——** subtext between them is a storyline all its own.

Even worse, unbeknownst to the driven newscaster, her cameraman is really an undercover agent for the FBI working to break up a complicated money laundering ring. As the newscaster covers her not-so-serious "lifestyle" stories, the hapless cameraman risks life and limb in outrageously comedic chase and fight scenes. As he pays less and less attention to the newscaster, she finds herself increasingly attracted to him. The romantic tension between the two of them is charming, and the contrast between the cameraman's boring film job and his secret FBI life is hysterical. The plot twists and turns make for a movie that is both hilarious and endearing.

- This movie has received excellent reviews from major, credible film critics, generally on the order of 3.5 out of 4 stars. As a combination romance/comedy, critics believe it will have wide appeal to both men and women.
- The film will open February 1, 2004. Though distributed nationally, this movie will be shown primarily where the **Black/——** population is most concentrated and in theaters where **Black/——** moviegoers are most likely to attend.
- *Large budget in control condition*: This movie is supported by a very **large/small** promotional budget of $27/$1.3 million. Ad placements in **national television and print outlets will dominate/will focus on local cable and outdoor outlets**—potential viewers **will be made well aware of this movie's release/may not be aware of this movie before it is released**.
- *Omitted in control condition*: **Veteran actors Angela Basset and Will Smith turn in outstanding performances** in this film as the newscaster and her cameraman./**Previously unknown actors Laquisha Brown and Tyrone Jefferson make their successful film debuts** in this movie as the newscaster and her cameraman.

Killer Mind. The movie begins with a clever twist, as a forensic investigator examines and dissects a grisly crime scene. It seems that people are being killed by a serial killer, and clues are few and far between. The local police team, led by a new detective and his veteran partner, avoid saying the killings were done by a serial killer until it becomes obvious and they absolutely have to—concern about public outcry weighs heavily on their minds, but they want to keep the details secret to avoid "copy cat" killings. When the new detective's boss wants to take him off of the case because he's new to the department, he successfully fights for his right to be the first **Black/——** in his department to work a serial murder. Issues of **race/officer loyalty** form a subtle backdrop against which the story unfolds, and the detectives are surprised to find out who their true friends are.

Eventually, a world-renowned psychiatrist who lectures on serial killers is brought in on the case. An academic recluse, she reluctantly joins the detectives in their search for the murderer. As she examines the forensic evidence collected in the opening scene and subsequent crime scenes, she realizes that the killer is her cousin. It turns out the killer is playing games with the psychiatrist, making his capture one of the scariest and most exciting scenes in the movie.

- Thought to be a "sleeper" in 2004, this action/suspense movie has garnered uniformly outstanding reviews from all critics, and it is believed that it will be in contention for several major awards. Though there are some graphic scenes, critics believe the combination of intrigue, mental gamesmanship, and clever dialog will draw both men and women into its sinister plot.

- The film will open January 30, 2004. Though distributed nationally, this movie will be shown primarily where the **Black/———** population is most concentrated and in theaters where **Black/———** moviegoers are most likely to attend.
- *Large budget in control condition*: This movie is supported by a very **large/small** promotional budget of $27/$1.3 million. Ad placements in **national television and print outlets will dominate/will focus on local cable and outdoor outlets**—potential viewers **will be made well aware of this movie's release/may not be aware of this movie before it is released.**
- *Omitted in control condition*: **Veteran actors Will Smith and Angela Basset turn in outstanding performances** in this film as the new detective and his partner./**Previously unknown actors Tyrone Jefferson and Laquisha Brown make their successful film debuts** in this movie as the new detective and his partner.

Appendix B
Diversity-Seeking Measure

To our knowledge, a diversity-seeking scale does not exist. Although the Miville-Guzman Universality Diversity Orientation (Miville et al. 1999) and Receptivity to Dissimilar Others (Douthitt, Eby, and Simon 1999) scales are close, we believe that they reflect passive tolerance or acceptance of diversity rather than the proactive seeking out of diversity that we believe drives crossover consumption behaviors. Therefore, we created our single-item measure, "I like seeing movies about other kinds of people," to capture this individual difference trait in a movie-viewing context. Although a multi-item scale would be preferred in general, research suggests that a multi-item measure is neither always desirable nor necessary (Drolet and Morrison 2001). Indeed, additional analyses suggest that our single item is more than adequate in its application in Studies 1 and 3 because it is related to other measures of diversity seeking.

Specifically, additional testing (n = 95) regarding the preliminary development of a diversity-seeking scale shows that this single item is correlated with a 12-item measure ($\alpha = .85$) of diversity seeking ($r = .27, p < .01$). Initial data analyses pertaining to the development of the diversity-seeking scale show that this scale comprises four subfactors (confirmatory factor analysis: $\chi^2 = 53.23, p = .28$; adjusted goodness-of-fit index = .87). Our single-item measure was significantly correlated with three of the four subscales listed in Table A1 (learning: $r = .27, p < .01$; hiring: $r = .20, p < .06$; eating: not significant; living: $r = .25, p < .01$), suggesting that different facets of general diversity-seeking tendencies have different influences on domain-specific consumption behaviors. In addition, this item was correlated with the number of diversity-related behaviors (e.g., "I have been to a same-sex wedding," and "I have spent the night in the home of someone of another ethnic group") out of 18 in which participants claimed to have engaged ($r = .23, p < .02$). Therefore, we are confident that this measure serves as an adequate representation of the diversity-seeking orientation that we believe moderates the impact of product ethnic orientation on attitudes and behavioral intentions.

REFERENCES

Aaker, Jennifer A., Anne M. Brumbaugh, and Sonya A. Grier (2000), "Non-Target Market Effects and Viewer Distinctiveness: The Impact of Target Marketing on Attitudes," *Journal of Consumer Psychology*, 9 (3), 127–40.

Acapulco Black Film Festival (1997), "State of Black Film," presentation at the Acapulco Black Film Festival, moderated by Warrington Hudlin, Acapulco, Mexico.

Ajzen, Icek and Martin Fishbein (1977), "Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research," *Psychological Bulletin*, 84 (5), 888–918.

Alexander, George (2000), "Fade to Black," *Black Enterprise*, (December), 107–115.

——— (2001), "Reaching the Silver Screen," *Black Enterprise*, (December), 89–93.

Andruss, Paula Lyon (2004), "'Dora' Translates Well," *Marketing News*, (October 13), 8.

Appiah, Osei (2001a), "Black, White, Hispanic, and Asian American Adolescents' Responses to Culturally Embedded Ads," *Howard Journal of Communications*, 12 (1), 29–48.

——— (2001b), "Ethnic Identification on Adolescents' Evaluations of Advertisements," *Journal of Advertising Research*, 41 (5), 7–22.

——— (2004), "It Must Be the Cues: Racial Differences in Adolescents' Responses to Culturally Embedded Ads," in *Diversity in Advertising*, Jerome D. Williams, Wei-Na Lee, and Curtis P. Haugvedt, eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 217–36.

Austin, Bruce A. (1989), *Immediate Seating: A Look at Movie Audiences*. Belmont, CA: Wadsworth.

Bannon, Lisa (2000), "Is the Country Ready for a Hispanic 'Rocky'? Hollywood Thinks Not," *The Wall Street Journal*, (April 6), A1.

Baron, Reuben M. and David A. Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (6), 1173–82.

Bearden, William O. and Michael J. Etzel (1982), "Reference Group Influence on Product and Brand Purchase Decisions," *Journal of Consumer Research*, 9 (3), 183–94.

Billig, Michael (1997), "Keeping the White Queen in Play," in *Off White: Readings on Race, Power, and Society*, Michelle Fine, Lois Weis, Linda C. Powell, and L. Mun Wong, eds. New York: Routledge, 149–66.

Brumbaugh, Anne M. (2002), "Source and Nonsource Cues in Advertising and Their Effects on the Activation of Cultural and Subcultural Knowledge on the Route to Persuasion," *Journal of Consumer Research*, 29 (2), 258–69.

Campos-Flores, Arian (2001), "Changing of the Guard; New Groups of Strivers Are Displacing the Old Ones," *Newsweek*, (August 6), 48.

Carter, Kelly (2001), "Studios' Views of Black Films May Undercut Success," *USA Today*, (April 11), D4.

Clunis, Andrew (2004), "No Sell-Out: Choice Bosses Deny Rumours," *The Voice*, (October 18–24), 8.

Cooper-Martin, Elizabeth (1991), "Consumers and Movies: Some Findings on Experiential Products," in *Advances in Consumer*

Research, Vol. 18, Rebecca Holman and Michael Solomon, eds. Provo, UT: Association for Consumer Research, 372–78.

Cote, Joseph A., James McCullough, and Michael Reilly (1985), "Effects of Unexpected Situations on Behavior-Intention Differences: A Garbology Analysis," *Journal of Consumer Research*, 12 (2), 188–94.

Davidson, A.R. and J.J. Jaccard (1979), "Variables That Moderate the Attitude-Behavior Relation: Results of a Longitudinal Survey," *Journal of Personality and Social Psychology*, 37 (8), 1364–76.

De Vany, Arthur S. and W. David Walls (1997), "The Market for Motion Pictures: Rank, Revenue, and Survival," *Economic Inquiry*, 35 (4), 783–94.

Devine, P.G. (1989), "Stereotypes and Prejudice: Their Automatic and Controlled Components," *Journal of Personality and Social Psychology*, 56 (1), 5–18.

Diamond, Helene (1989), "Lights, Camera, Action! A Film's Not a Wrap Until Test Audience Says So," *Marketing News*, 23 (19), 10–11.

Douthitt, Shane S., Lillian T. Eby, and Shana A. Simon (1999), "Diversity of Life Experiences: The Development and Validation of a Biographical Measure of Receptiveness to Dissimilar Others," *International Journal of Selection and Assessment*, 7 (2), 112–25

Drolet, Aimee L. and Donald G. Morrison (2001), "Do We Really Need Multiple-Item Measures in Service Research," *Journal of Service Research*, 3 (February), 196–204.

Easton, Nina J. (1991), "Good News/Bad News of the New Black Cinema; The Films Are Finally Coming Out in Force, but Marketing African-American Themes to a Broad Audience Is a Challenge Now," *Los Angeles Times*, (June 16), 5.

Fisher, Robert J. and Linda L. Price (1992), "An Investigation into the Social Context of Early Adoption Behavior," *Journal of Consumer Research*, 19 (4), 477–86.

Forehand, Mark R. and Rohit Deshpandé (2001), "What We See Makes Us Who We Are: Priming Ethnic Self-Awareness and Advertising Response," *Journal of Marketing Research*, 38 (August), 336–48.

Friedman, Wayne, Jon Fine, Tobi Elkin, and Ann Marie Kerwin (2003), "Print Feels Heat as Movie Dollars Shift," *Advertising Age*, (September 1), 17.

Fuertes, Jairo N., Marie L. Miville, Jonathan J. Mohr, William E. Sedlacek, and Denise Gretchen (2000), "Factor Structure and Short-Form of the Miville Guzman Universality-Diversity Scale," *Measurement and Evaluation in Counseling and Development*, 33 (3), 157–69.

Garafalo, Reebee (1993), "Black Popular Music: Crossing Over or Going Under," in *Popular Music: Politics, Policies, and Institutions*, Tony Bennett, Simon Frith, Lawrence Grossberg, John Shepherd, and Graeme Turner, eds. London: Routledge, 231–48.

——— (1994), "Culture Versus Commerce: The Marketing of Black Popular Music," *Public Culture*, 7 (1), 275–87.

Gerbner, George, Larry Gross, Michael Morgan, and Nancy Signorielli (1994), "Growing Up with Television: The Cultivation Perspective," in *Media Effects: Advances in Theory and Research*, Jennings Bryant and Dolf Zillmann, eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 17–41.

Gibbs, James L. (1999), "Television and Diversity: A Quantum Leap Model," in *A Nation Divided: Diversity, Inequality, and Community in American Society*, Phyllis Moen, Donna Dempster-McLain, and Henry Walker, eds. Ithaca, NY: Cornell University Press, 260–74.

Grayson, Kent and Radan Martinec (2004), "Consumer Perceptions of Iconicity and Indexicality and Their Influence on Assessment of Authentic Market Offerings," *Journal of Consumer Research*, 31 (2), 296–312.

Green, Corliss L. (1999), "Ethnic Evaluations of Advertising: Interaction Effects of Strength of Ethnic Identification, Media Placement, and Degree of Racial Composition," *Journal of Advertising*, 28 (1), 49–64.

Grier, Sonya A. and Anne M. Brumbaugh (1999), "Noticing Cultural Differences: Ad Meanings Created by Target and Non-Target Markets," *Journal of Advertising*, 28 (Spring), 79–93.

——— and Rohit Deshpandé (2001), "Social Dimensions of Consumer Distinctiveness: The Influence of Social Status on Group Identity and Advertising Persuasion," *Journal of Marketing Research*, 38 (May), 216–24.

Guess, Terri P. (2004), "Diversity at the Fancy Food Show; Black Exhibitors Applaud New Pavilion for Minorities," *Network Journal*, 10 (8), 25.

Halter, Marilyn (2000), *Shopping for Identity: The Marketing of Ethnicity*. New York: Schocken Books.

Hirschman, Elizabeth C. and Morris B. Holbrook (1982), "Hedonic Consumption: Emerging Concepts, Methods, and Propositions," *Journal of Marketing*, 46 (July), 92–101.

Hoyer, Wayne D. and Steven P. Brown (1990), "Effects of Brand Awareness on Choice for a Common, Repeat-Purchase Product," *Journal of Consumer Research*, 18 (2), 141–48.

Hughes, Alan (2002), "Hip Hop Economy," *Black Enterprise*, (May), 70–75.

Johnson, Lynn D. (2003), "Independent Black Filmmakers on Hollywood: The Distribution of Black Film," *Bright Lights Film Journal*, (accessed July 14, 2003), [available at www.brightlightsfilm.com/36/distribution1.html].

Kahn, Barbara E. (1993), "The Influence of Positive Affect on Variety Seeking Among Safe, Enjoyable Products," *Journal of Consumer Research*, 20 (2), 257–70.

——— (1995), "Consumer Variety-Seeking Among Goods and Services," *Journal of Retailing and Consumer Services*, 2 (3), 139–48.

Kozup, John C., Elizabeth H. Creyer, and Scot Burton (2003), "Making Healthful Food Choices: The Influence of Health Claims and Nutrition Information on Consumers' Evaluations of Packaged Food Products and Restaurant Menu Items," *Journal of Marketing*, 67 (April), 19–34.

Kraus, Stephen J. (1995), "Attitudes and the Prediction of Behavior: A Meta-Analysis of the Empirical Literature," *Personality and Social Psychology Bulletin*, 21 (1), 58–75.

Laroche, Michel, Chankon Kim, and Lianxi Zhou (1996), "Familiarity and Confidence as Determinants of Purchase Intention: An Empirical Test in a Multiple Brand Context," *Journal of Business Research*, 37 (2), 115–20.

Loken, Barbara and James Ward (1990), "Alternative Approaches to Understanding the Determinants of Typicality," *Journal of Consumer Research*, 17 (2), 111–16.

Martin, Ingrit M. and David W. Stewart (2001), "The Differential Impact of Goal Congruency on Attitudes, Intentions, and the Transfer of Brand Equity," *Journal of Marketing Research*, 38 (November), 471–84.

Martin, Leonard L., John J. Seta, and Rick A. Crelia (1990), "Assimilation and Contrast as a Function of People's Willingness and Ability to Expend Effort in Forming an Impression," *Journal of Personality and Social Psychology*, 59 (1), 27–37.

Masters, Kim (1996), "Testing the Faith: Preachers White Shakeup Hollywood's Belief That Whites Won't Go to Black Movies," *Time*, (December 16), 73–74.

McGuire, William J., Claire V. McGuire, Pamela Child, and Terry Fujioka (1978), "Salience of Ethnicity in the Spontaneous Self-Concept as a Function of One's Ethnic Distinctiveness in the Social Environment," *Journal of Personality and Social Psychology*, 36 (5), 511–20.

Mellgren, James (2001), "The Bold New World of Condiments," *Gourmet Retailer*, (October 1), 132–35.

Miville, Marie L., Charles Gelso, Raji J. Pannu, Will Liu, Pegah Touradji, Pauline Holloway, and Jairo Fuertes (1999), "Appreciating Similarities and Valuing Differences: The Miville-

Guzman Universality-Diversity Scale," *Journal of Counseling Psychology*, 46 (3), 291–307.

Monteith, M.J. (1993), "Self-Regulation of Prejudiced Responses: Implications for Progress in Prejudice Reduction Efforts," *Journal of Personality and Social Psychology*, 65 (3), 469–85.

Pampel, Fred, Dan Fost, and Sharon O'Malley (1994), "Marketing the Movies," *American Demographics*, 16 (3), 48–52.

Quester, Pascale G. and Justin Smart (1998), "The Influence of Consumption Situation and Product Involvement over Consumers' Use of Product Attribute," *Journal of Consumer Marketing*, 15 (3), 220–38.

Ratner, Rebecca K. and Barbara E. Kahn (2002), "The Impact of Private Versus Public Consumption on Variety-Seeking Behavior," *Journal of Consumer Research*, 29 (2), 246–57.

Ray, Paul H. and Sherry H. Anderson (2000), *The Cultural Creatives: How 50 Million People Are Changing the World*. New York: Harmony Books.

Retail World (2000), "Asian Food Sales Double," (August 7–18), 16.

Richins, Marsha L. (1994), "Valuing Things: The Public and Private Meanings of Possessions," *Journal of Consumer Research*, 21 (December), 504–521.

Sheppard, Blair H., Jon Hartwick, and Paul R. Warshaw (1988), "The Theory of Reasoned Action: A Meta-Analysis of Past Research," *Journal of Consumer Research*, 15 (3), 325–43.

Stayman, Douglas M. and Rohit Deshpandé (1989), "Situational Ethnicity and Consumer Behavior," *Journal of Consumer Research*, 16 (3), 361–71.

Strong, Andrea (2002), "Worldly Sweets," *Restaurant Business*, (March 15), 66–71.

Tharp, Marye C. (2001), *Marketing and Consumer Identity in Multiethnic America*. Thousand Oaks, CA: Sage Publications.

Tversky, Amos and Daniel Kahnemann (1974), "Judgment Under Uncertainty: Heuristics and Biases," *Science*, (September 27), 1124–31.

Wakefield, Kirk L. and J. Jeffrey Inman (2003), "Situational Price Sensitivity: The Role of Consumption Occasion, Social Context, and Income," *Journal of Retailing*, 79 (4), 199–212.

Walkup, Carolyn (2003), "Mainstream Eateries Catch Sushi Fever, Net Raw Fish for Menus," *Nation's Restaurant News*, (April 7), 8–9.

Watts, E.K. and M. Orbe (2002), "The Spectacular Consumption of 'True' African American Culture: 'Whassup' with the Budweiser Guys?" *Critical Studies in Media Communication*, 19 (1), 1–20.

Whittler, Tommy E. (1989), "Viewers' Processing of Actor's Race and Message Claims in Advertising Stimuli," *Psychology and Marketing*, 6 (4), 287–309.

——— (1991), "The Effects of Actors' Race in Commercial Advertising: Review and Extension," *Journal of Advertising*, 20 (1), 54–60.

——— and Joan DiMeo (1991), "Viewers' Reactions to Racial Cues in Advertising Stimuli," *Journal of Advertising Research*, 31 (6), 37–46.

Williams, Jacqueline A. (1995), "The Effect of Ethnic and American Identification on Consumers' Responses to Ethnic Advertising Appeals," doctoral dissertation, Department of Marketing, Florida State University.

Williams, Jerome D., William J. Qualls, and Sonya A. Grier (1995), "Racially Exclusive Real Estate Advertising: Public Policy Implications for Fair Housing Practices," *Journal of Public Policy & Marketing*, 14 (Fall), 225–44.

Wolf, Michael J. (1999), *The Entertainment Economy: How Mega-Media Forces Are Transforming Our Lives*. New York: Random House.

Xu, Jing, Soyeon Shim, Sherry Lotz, and David Almeida (2004), "Ethnic Identity, Socialization Factors, and Culture-Specific Consumption Behavior," *Psychology and Marketing*, 21 (2), 93–112.

Yi, Daniel and Ray Herndon (2001), "Diversity Turns O.C. into Mixer; Demographics: Census Data Show Increasing Chance of Encountering Those of Different Race or Ethnic Background," *Los Angeles Times*, (May 29), B1.

Zalesny, Mary D. and J. Kevin Ford (1990), "Extending the Social Information Processing Perspective: New Links to Attitudes, Behaviors, and Perceptions," *Organizational Behavior and Human Decision Processes*, 47 (2), 205–246.

Chris Pullig, Carolyn J. Simmons, & Richard G. Netemeyer

Brand Dilution: When Do New Brands Hurt Existing Brands?

Potential for trademark dilution occurs when a new brand introduces itself with an identical or substantially similar mark to that of an existing brand. Currently, there is neither a clear standard for legal proof of dilution nor widely accepted measures for managerial use. Using a shared-brand-network model, the authors use three different but conceptually related measurement methodologies (response latency, aided recall, and simulated choice) across five experiments to examine "blurring" dilution effects. The results suggest that when junior brands operate in similar categories and position on similar attributes, there is no immediate threat of blurring dilution. However, if the junior brand chooses to position on dissimilar attributes in a similar category, dilution of attribute associations is likely. When the junior brand operates in dissimilar categories, there is a threat of blurring dilution for both category and attribute associations. The final two studies demonstrate that dilution of junior brands (i.e., operating in dissimilar categories positioned on dissimilar attributes) can suppress consideration and choice probabilities for the senior brand and that this effect may extend over a period of time.

Recently, the United States Supreme Court unanimously ruled in favor of a small store against a world-famous lingerie manufacturer and retailer (*Mosely v. V. Secret Catalogue* 2003). The owners of the small store, which sells lingerie and other "adult novelty" items, named their store Victor's Secret. Shortly after opening, they were sued by Victoria's Secret, which claimed protection of its brand name under the Federal Trademark Dilution Act of 1995. At issue before the Supreme Court was the legal standard for proof in trademark dilution cases. In its decision, the Court ruled that trademark holders (i.e., the *senior* brand, Victoria's Secret) must show evidence of "actual dilution" by the *junior* brand (i.e., Victor's Secret). Showing a mental association between the two marks or a likelihood of dilution is insufficient. At the same time, however, trademark holders do not need to prove actual lost sales to show that dilution has occurred.

These vague guidelines define a broad area within which case law establishes what constitutes the legal standard of proof for trademark dilution. However, how to establish and/or measure dilution is not only a legal and public policy issue but also one of concern to brand managers who work to maintain and leverage brand equity. This article addresses two questions pertinent to this issue: (1) How can and should trademark dilution be measured? and (2) In what situations is dilution likely to occur in response to the presence of a junior brand?

Background

Trademark Infringement Versus Dilution

Trademark dilution differs from the more familiar legal doctrine of trademark infringement. "Trademark infringement" occurs when a competing party (i.e., a junior brand) uses an identical or substantially similar mark (e.g., brand name, slogan, symbol) that is already being used by an existing party (i.e., a senior brand) such that consumers are likely to be confused, mistaken, or deceived about the source of the goods being sold. However, infringement can extend beyond competing brands. If a "not insubstantial" proportion of consumers are confused because they think a brand has been extended into an unrelated category, a finding of infringement would also be likely (Jacoby 2001). Thus, as a legal doctrine, the primary focus of trademark infringement is the protection of consumers from misleading information. Conversely, "trademark dilution" is defined as "the lessening of the capacity of a famous mark to identify and distinguish goods or services, *regardless of the presence or absence of* (1) competition between the owner of the famous mark and other parties or (2) likelihood of *confusion, mistake, or deception*" (Federal Trademark Dilution Act 1995, § 4, emphasis added). Thus, the primary focus of dilution is the protection of the brand asset, or the protection of brand equity (Feldwick 1996; Keller 1993).

Dilution may take two broad forms: (1) tarnishment and (2) blurring (Simonson 1993). "Tarnishment" refers to a lowered evaluation of a senior brand due to a junior brand. The damage of tarnishment is fairly obvious in that there is the attachment of a negative association to the senior brand. For example, consider the case in which the American

---

Chris Pullig is Assistant Professor of Marketing, Hankamer School of Business, Baylor University (e-mail: Chris_Pullig@Baylor.edu). Carolyn J. Simmons is Visiting Associate Professor of Marketing, Williams School of Commerce, Economics, and Politics, Washington and Lee University (e-mail: simmonsc@wlu.edu). Richard G. Netemeyer is Ralph A. Beaton Chair of Free Enterprise, McIntire School of Commerce, University of Virginia (e-mail: rgn3p@virginia.edu). This research was funded in part by research grants from the Hankamer School of Business at Baylor University, Pamplin College of Business at Virginia Polytechnic Institute and State University, and McIntire School of Commerce at the University of Virginia.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Express slogan "Don't leave home without it" was used to promote condoms by another firm (*American Express Co. v. Vibra Approved Lab. Corp.* 1989). American Express believed that the association of condoms with its slogan could create a negative association for its traveler's check brand, even though there was likely no confusion that it was the maker of the condoms.

"Blurring" has been described as the "gradual whittling away" of the identity of the mark (brand) in the mind of the public (Schechter [1927] 1970) and/or the lessening of the uniqueness of the brand in the minds of consumers (Simonson 1993) due to the emergence of a junior brand. Both of these perspectives hint at the weakening of important brand associations (e.g., product category, specific attributes). This weakening of associations may affect brand choice by making consumers less likely to consider the brand when making a decision in the category and less likely to associate it confidently with desired attributes (Alba, Hutchinson, and Lynch 1991). For example, the Lexis data service argued that the introduction of the Lexus automobile potentially added a new category association that could blur the meaning of the brand in consumers' minds (*Mead Data Central Inc. v. Toyota Motor Sales, USA Inc.* 1989). Although both forms of dilution merit examination, blurring is of more interest because it is less obvious, is less frequently studied, and lacks consensus on how it should be assessed. Thus, this article focuses on blurring dilution.

Likelihood of Dilution or Actual Dilution

A thorny issue in trademark dilution is how to prove dilution. Before the Supreme Court's ruling in the *Victoria's Secret* case, a split of interpretation had developed between the Second and Fourth Circuits about whether the plaintiff in a trademark dilution case needed to prove actual harm to its famous mark or a mere likelihood of harm (Lee 2004; McCarthy 2004). A superficial reading of the ruling suggests that the Court took the Fourth Circuit's side in requiring proof of actual dilution. However, the Court indicated that proof of actual dilution does not require evidence of lost sales or profits. Furthermore, it added that dilution could be shown through circumstantial evidence, such as a mental association between the junior and the senior marks, when the two marks are completely identical. This part of the ruling has created substantial confusion by suggesting that proof of likelihood of dilution *can* be adequate.

McCarthy (2004) interprets this ruling as creating two types of cases: one in which the marks are identical and one in which they are not. In the case of identical marks, a mere mental association may be adequate circumstantial evidence of dilution. When the marks are not identical, proof of actual dilution is required. Other legal scholars (e.g., Lee 2004) view this part of the opinion as simply giving evidentiary weight to the degree of similarity between the two marks. They argue that it defies logic to apply two completely different evidentiary standards to a mark that is identical in every aspect compared with a mark that is only slightly different (e.g., varies in font style).

Furthermore, subsequent decisions have applied the evidentiary standard differently. Some district courts have interpreted the opinion to mean that when competing marks are *completely identical*, mere mental association (i.e., likelihood of dilution) is sufficient circumstantial evidence of dilution. However, these decisions find that there is dilution only after a finding of likelihood of confusion, that is, trademark infringement (e.g., *Nike v. Variety Wholesalers Inc.* 2003). Other district courts have interpreted the Court's ruling to mean that a mental association is only one piece of evidence and that additional evidence of actual dilution is required even when the marks are identical (e.g., *Savin Corp v. The Savin Group* 2003).

In summary, this split in legal opinion has generated considerable debate about how to demonstrate and measure dilution. Thus, a better understanding of how dilution can be measured can contribute to this debate and offer insights for managers who are concerned about maintaining brand equity and defending against trademark dilution.

Dilution as a Change in Brand Knowledge

Implicit in the legal doctrine of trademark dilution is the notion that dilution is a change in brand knowledge (Jacoby 2001). In the case of blurring, it is the weakening of brand associations, such as the brand's category, attributes, advertising tag lines, and so forth. This view is consistent with the suggestion in brand management and trademark dilution literature that a brand's value or equity resides in consumers' brand knowledge or memory network.

For example, Keller (1993, p. 2) defines customer-based brand equity as "the differential effect of brand knowledge on consumer response to the marketing of a brand." Two prior studies also view trademark dilution as a change in brand knowledge. Peterson, Smith, and Zerrillo (1999) focus on "brand dominance" and "brand typicality," which refer to the probability that a brand will be recalled given its category as a retrieval cue and the probability that a category will be recalled given the brand name as a retrieval cue, respectively. Morrin and Jacoby's (2000) conceptualization is broader; it considers the weakening of associations between the brand and any of its aspects or associations, not just its category. Finally, legal scholars and public policy researchers suggest that dilution by blurring involves the formation of a shared brand knowledge network, in which knowledge about the senior and junior brands may compete for activation when a consumer thinks about the senior brand (Jacoby 2001; McCarthy 2004). Jacoby (2001, p. 1048) notes that the harm of dilution could take several forms:

> At the very least, the consumer would be required to engage in additional cognitive effort to parse out which network applied to the current circumstances [senior or junior brand]. Generally, the more effortful the activity (including a mental activity), the less likely it is to be performed.... "[L]aboratory studies reveal that ... trademark dilution can reduce the strength of pre-existing brand associations through the creation of additional nodes in consumer's brand based memory networks. Further, exposure to the second comer's [junior brand] promotional use of the mark can reduce the accuracy and, to some extent, the speed of retrieval from memory of the first-user [senior brand] information.

Our conceptualization of dilution is consistent with these writings.[1]

The Conceptualization and Measurement of Dilution

We view blurring as a weakening of the associations between a brand and its distinctive aspects (e.g., category, attributes, benefits) that define the meaning of the brand in consumers' minds. We now consider how brand associations are used in decision making and derive measures of dilution based on this conceptualization.

When consumers retrieve brand knowledge from memory, they may (1) use desired brand aspects as retrieval cues for a brand name or (2) use the brand name as a retrieval cue to access knowledge of the brand's aspects. Thus, we focus on the associative strength underlying each of these two processes. "Brand name accessibility" refers to the likelihood that the brand will come to mind when a brand aspect is used as a retrieval cue (e.g., how likely it is that Big Red will come to mind when a person thinks of chewing gum or cinnamon flavor). When brand name accessibility decreases (i.e., is diluted), it is less likely that the brand will be recalled and enter the consumer's consideration set in decision-making situations (Alba, Hutchinson, and Lynch 1991). "Aspect accessibility" refers to the likelihood that a given brand aspect will come to mind when the brand name is used as a retrieval cue (e.g., how likely it is that chewing gum or cinnamon flavor will come to mind when a person thinks about the Big Red brand). When aspect accessibility decreases (i.e., is diluted), even consumers who seek the benefits offered by the brand are less likely to consider and choose the brand (Alba, Hutchinson, and Lynch 1991).[2]

On the basis of this conceptualization, we use three dilution measurement methods: (1) response latency to recognize brand associations, (2) recall of brand associations, and (3) brand choice. The first two measures directly assess brand name and aspect accessibility. Although both have been advocated as useful in the assessment of dilution (Morrin and Jacoby 2000; Peterson, Smith, and Zerrillo 1999; Simonson 1993), the response latency measure has been singled out as particularly relevant. Reduction in the speed at which people are able to access, recognize, and verify brand associations is what Jacoby (2001, p. 1049) refers to as "the essence of a weakening of associations" for dilution. In contrast, our third measure addresses the consequences of reduced brand name or aspect accessibility, namely, reduced probability of consideration and brand choice.

The Shared Brand Network and Predicting Dilution

To understand how a junior-brand strategy may affect the senior brand, we consider how brand information is stored in and retrieved from memory. Associative network theory (Anderson 1983) suggests that brand information is encoded in long-term memory as a pattern of linkages between concept nodes (e.g., associations between the brand and its aspects). Although the senior and junior brands are not related from the perspectives of ownership, production, or marketing, they are related by a similar mark (e.g., name, logo, defining slogan). Thus, although consumers are likely to possess cognitive networks for both the senior and the junior brands, there is some connection between these networks due to the shared/similar trademark. To the extent that there are other similarities between the two brands (i.e., similar aspects), these two networks become even more interconnected (Jacoby 2001). Recent memory research shows that such related information is often encoded in one composite memory trace (Humphreys et al. 2000; Shiffrin 2003). However, regardless of whether the junior brand shares a composite memory trace with the senior brand, it is linked to the senior brand in memory through what we call a "shared brand network."

What are the implications of this shared brand network for the brand name and aspect accessibility of the senior brand? Consider how consumers retrieve knowledge from memory. At retrieval, cues activate corresponding nodes in memory, and this activation spreads to related constructs. The speed at which activation spreads is determined by the strength and proximity of linkages among constructs. The closer constructs are encoded in memory, and the stronger the linkages among them, the greater is the likelihood that related constructs will be activated. Because the senior and junior brands reside in a shared brand network, thinking about the senior brand activates associations for the junior brand (Anderson, Green, and McCulloch 2000). Furthermore, because all associated constructs compete for activation, these junior-brand associations are likely to reduce the probability and speed of activation of senior-brand associations (Burke and Srull 1988). However, this competitive interference is likely moderated by the degree of similarity between the aspects of the senior and junior brands. Indeed, adding similar constructs to the shared brand network (e.g., when the junior brand shares similar aspects with the senior brand) is likely to increase the probability and speed of activation of senior-brand associations through the spreading of activation among these closely related constructs (Humphreys et al. 2000).

---

[1] Peterson, Smith, and Zerrillo (1999) and Morrin and Jacoby (2000) are the first to measure and examine moderators of trademark dilution. Their studies have many strengths, such as an empirical investigation of brand-knowledge-based measures of dilution, the examination of multiple brands, the use of recognition and recall measures, and the use of brand names and logos as stimuli. Nonetheless, their studies suggest the need for further empirical work. Limitations include weak and inconsistent effects, effects that occur only in comparison to a condition in which the senior brand's associations are reinforced, pooling of effects so that it is not possible to examine differential effects for many specific brand associations, and a limited examination of a key moderator of dilution (attribute similarity is not examined). We extend their pioneering work with a research design that can address all these issues.

[2] Keller (1993) divides brand knowledge into two similar components: brand awareness and brand image. Our brand name accessibility component parallels his brand awareness component, but our aspect accessibility component is less inclusive than his brand image component, because we include only specific associations and not brand attitudes. Our focus on specific associations follows from understanding blurring as a weakening of association between a brand and its distinctive aspects.

Therefore, we expect that when the junior brand's aspects (e.g., category, attributes) are dissimilar to the senior brand, the senior brand's name accessibility and aspect accessibility decrease; that is, the senior brand experiences blurring dilution. When the junior brand's aspects are similar to those of the senior brand, the senior brand's name accessibility and aspect accessibility increase; that is, the senior brand experiences reinforcement of its distinctive associations. Although somewhat similar effects have been found in the brand extension literature, there are two key differences between this stream of research and our studies: First, most research in extension dilution examines inconsistent or failed extensions, relying on measures of global brand evaluations (cf. John, Loken, and Joiner 1998), not accessibility (cf. Morrin 1999). Second, brand extensions are known to be related to the parent brand. In our research, the senior and junior brands are known to be unrelated. Although dilution applies equally to situations in which consumers are or are not confused about the relationship between the junior and the senior brands (Jacoby 2001), dilution comes to the forefront only when there is no confusion.

How Do These Effects Occur In the Absence of Confusion?

It may seem implausible that the dilution and reinforcement effects we predict occur without confusion about the relationship between the senior and the junior brands. Why would the consumer not simply use his or her knowledge that the brands are unrelated to selectively "ignore" junior-brand associations when thinking about the senior brand? Again, consider how knowledge is retrieved from memory. When a stimulus such as a brand name is encountered, its associations are activated according to their accessibility. Highly accessible associations contribute to an "automatic interpretation" of the construct. This automatic interpretation incorporates any blurring or reinforcement that results from associations with the junior brand and its aspects in the shared brand network. This automatic interpretation may be corrected through the consideration of less accessible associations or externally available information, such as the knowledge that the brands are unrelated. However, such correction requires cognitive effort, and as a result, undercorrection is common (Jacoby 2001; Johar and Simmons 2000). Thus, the accessibility of senior-brand associations is influenced by junior-brand associations, even when there is no confusion, and dilution or reinforcement is the consequence.

Study 1: Effects of Product Category Similarity on Response Latencies

The effect of category similarity on dilution is an important issue because the trademark dilution doctrine extends to noncompetitive brands (i.e., brands in different categories) and to brands in direct competition with the senior brand. Understanding where the danger lies can help public policy makers and brand managers monitor and control dilution. We examine the moderating effects of category similarity on reinforcement and dilution of the senior brand while holding the similarity of other brand aspects constant at a relatively neutral level. Recall that we expect high category similarity to lead to reinforcement of the senior brand and low category similarity to lead to dilution of the senior brand. We focus on the case in which the junior brand's name is identical to that of the senior brand.

H_1: Category similarity moderates the effect of a same-named junior brand on the senior brand's distinctive associations.

a. When category similarity is high, the senior brand is reinforced (i.e., brand name accessibility and aspect accessibility increase).

b. When category similarity is low, the senior brand is diluted (i.e., brand name accessibility and aspect accessibility decrease).

Method

Overview. Figure 1 displays our procedure and analyses. Our senior brand is Big Red, a chewing gum that is strongly associated with cinnamon flavor. In Step 1, we expose participants to a series of advertisements, including one for a new brand (the junior brand or an identical unbranded product). The new product belongs either to a category that is highly similar to that of the senior brand (bubble gum) or to one that is dissimilar to that of the senior brand (snack bars). In Step 2, we measure two dependent variables: (1) senior–brand name accessibility, or the response latency to recognize correctly the brand name as "going with" the distinctive aspect given as a cue, and (2) senior-brand aspect accessibility, or the response latency to recognize correctly a distinctive aspect as "going with" the brand name given as a cue. (The senior brand's distinctive aspects are chewing gum and cinnamon flavor.) In Step 3, we determine whether exposure to the junior brand increases or decreases these accessibilities compared with exposure to an identical unbranded product. If exposure to the junior brand increases brand name or aspect accessibility (i.e., if response latencies decrease), the senior brand has been reinforced by the junior brand. If exposure to the junior brand decreases brand name or aspect accessibilities (i.e., if response latencies increase), the senior brand has been diluted by the junior brand. Thus, our category similarity manipulation enables us to determine whether category similarity moderates reinforcement and dilution.[3]

Design and sample. Pretest participants indicated (on seven-point scales) that they were highly familiar with Big Red chewing gum (M = 6.73) and strongly associated Big Red with cinnamon flavor (M = 6.52). Thus, we focus on these two distinctive aspects of Big Red (i.e., chewing gum and cinnamon flavor). Pretest participants also identified bubble gum and snack bars as high (M = 5.26) and low (M = 2.05), respectively, in similarity to chewing gum; this difference is significant ($F_{1,18} = 42.39$, $p < .01$). On the basis of these results, our design is a 2 (category similarity:

---

[3] In Studies 1–3, we use an identical unbranded product as a control to isolate the effect of a junior brand. In Studies 4–5, we use a well-known brand as a control to increase ecological validity (consumers rarely encounter unbranded products) and external validity (our pattern of results is unaffected, increasing our confidence in their generality).

FIGURE 1

Procedure for Detecting Dilution of the Big Red → Chewing Gum Association

Dilution is shown if exposure to a junior brand increases the RT to recognize chewing gum as an association of the Big Red brand.

	1. Show Advertisements.			**2. Measure RT for the Association: Does chewing gum "go with" Big Red?**
Condition	**Ad 1 (Filler)**	**Ad 2 (Focal Advertisement)**	**Ad 3 (Filler)**	
High Similarity				
Unbranded	Skittles candy	Brand A bubble gum	Nutri-Grain snack bars	RT if not exposed to junior brand
Junior brand	Skittles candy	Big Red bubble gum	Nutri-Grain snack bars	– RT if exposed to junior brand
				RT difference
Low Similarity				
Unbranded	Skittles candy	Brand A snack bars	Bazooka bubble gum	RT if not exposed to junior brand
Junior brand	Skittles candy	Big Red snack bars	Bazooka bubble gum	– RT if exposed to junior brand
				RT difference

3. Examine how the RT for those exposed to the junior brand differs from the RT for those not exposed to the junior brand.

Dilution is shown by a negative difference. Reinforcement is shown by a positive difference.

Notes: RT = response time.

high [bubble gum] and low [snack bars]) × 2 (new product: junior brand [Big Red] and unbranded [Brand A]) × 2 (order of ad presentation) between-subjects design. We hold attribute similarity constant at a neutral level, as we describe subsequently. Participants were 89 randomly assigned undergraduate students.

Procedure and materials. In a computer-administered procedure, participants saw three advertisements: one for bubble gum, one for bite-size candy, and one for snack bars. Two of these were filler advertisements, which were intended to reduce hypothesis guessing. The third advertisement was for the new product, either a bubble gum (high category similarity) or a snack bar (low category similarity). Advertisements in each similarity condition were identical except for brand name (Big Red or Brand A), and they contained a disclaimer that stated that the product is not manufactured by the makers of Big Red (Brand A) chewing gum. We warned participants that some of the advertisements might be for products that were not yet branded and that these products would be identified with a letter. In all conditions, the two filler advertisements were for well-known products in the remaining two categories for that condition (Skittles bite-size candy and either Bazooka bubble gum or Nutri-Grain snack bars).

Each advertisement contained a brand name, a product category, and two claims. We told participants that the advertisements were depicted in this simple fashion because we were interested in their reactions to the informational content of the advertisements. In the new product advertisements, we held attribute similarity with Big Red chewing gum constant at a neutral level with one similar and one dissimilar attribute. For the bubble gum, claims were "great cinnamon flavor" (similar to Big Red chewing gum) and "blow awesome bubbles" (dissimilar). For the snack bar, the claims were "great cinnamon flavor" (similar) and "enriched with vitamins and minerals" (dissimilar). In a pretest, these two dissimilar attributes were not rated as differing in dissimilarity to Big Red chewing gum (M = 1.89 and 1.31; $F_{1, 18} = 3.31$, $p > .09$). Each advertisement was displayed for 30 seconds.[4] The new-product advertisement was always second in the sequence, and presentation of the filler advertisements was counterbalanced.

After presentation of the advertisements, we collected accessibility measures using procedures that Fazio (1990) described. First, we presented a brand name or an aspect. Participants focused on this cue and pressed the space bar when they were ready to see the second cue (an aspect or brand name, depending on which was not presented first). The second cue was then displayed on the next screen. Participants pressed a key to indicate whether the second cue went with the first cue (yes or no). They were instructed to work as quickly as possible without sacrificing accuracy. The computer captured the response latency (in milliseconds) from the time the second cue was displayed until a response key was pressed. Accessibility trails were preceded by a set of practice trials that were unrelated to the brands presented in the advertisements. Participants practiced until they were familiar with the procedure and could work quickly. The practice trials also provided a buffer between presentation of the advertisements and collection of dependent variables of interest. Following the accessibility trials, we administered manipulation checks for category similarity and lack of confusion between the senior and the junior brands. We also measured familiarity with Big Red.[5]

Results

Manipulation and data checks. Participants indicated (on seven-point scales) that they were highly familiar with Big Red chewing gum (M = 6.03) and strongly associated Big Red with cinnamon flavor (M = 6.38); thus, the chewing gum category and cinnamon flavor are good exemplars of the brand's distinctive aspects. As for the dissimilar attributes used in the advertisements for new products, participants indicated a weak association of Big Red with awesome bubbles (M = 1.92) and with vitamins and minerals (M = 1.19). They also identified bubble gum as high (M = 6.63) and snack bars as low (M = 2.95) in similarity to chewing gum; this difference is significant ($F_{1, 72} = 27.42$, $p < .001$). We also collected a seven-point scale measurement of liking for the Big Red brand. Liking the brand did not differ across conditions, indicating that our manipulations did not create tarnishment effects ($F_{3, 69} = 1.07$, $p > .10$). Finally, a free-recall measure of ad content showed that participants were not confused about the relationship between the senior brand and the junior brand; only 3.6% of participants did not report the disclaimer, and this number did not vary by condition ($\chi^2_3 = .19$, $p > .10$). We eliminated respondents who did not report the disclaimer.

Although we instructed participants to work as quickly as possible without sacrificing accuracy, it is possible that some traded off accuracy for speed. That is, effects on accessibilities could be reflected in the accuracy of responses rather than in response latencies. For example, participants could make errors such that dilution shows up as a lower probability of accurately identifying cinnamon as an aspect of Big Red rather than as a longer response latency to recognize this association. Thus, we checked our data for such errors. The average error rate was low (4.2%) and did not vary by condition (all $\chi^2 < 1.0$, $ps > .10$). This indicates that participant error in responding is not problematic. Still, we eliminated these errors from our analysis.

Hypothesis tests. We log-transformed response latencies to achieve distribution normality (Fazio 1990). We used each respondent's mean latency on filler trials as a covariate to isolate variance due to individual differences, thus reducing error variance and increasing the power of our statistical tests. Order of presentation of the filler advertisements had no impact on response latencies (*p*s > .10), and so we do not discuss its effects further.

---

[4] The 30-second display was intended to encourage participants to take the time to read the disclaimer. The viewing situation was more naturalistic (display time was not controlled) in Studies 3, 4, and 5. The results were similar.

[5] There were 92 trials in three blocks, with a 30-second rest between blocks. Experimental trials measured the accessibilities that we report in Table 1. Filler trials served to (1) decrease hypothesis guessing, (2) create a set of trials in which correct yes and no responses are equally represented and thus avoid creating a response bias, and (3) provide an average response time for each participant for use as a covariate. (The covariate isolates variance due to individual differences; this analysis should increase the sensitivity of our tests compared with Morrin and Jacoby's [2000] use of similar measures.) Within each block, we randomized the order of trials for each participant.

H_1 predicts that category similarity moderates the effect of the junior brand on the senior brand, such that the senior brand's association with its distinctive aspects is reinforced when category similarity is high (H_{1a}) and diluted when category similarity is low (H_{1b}). This prediction implies an interaction between category similarity and new product of a particular form. When similarity is high, senior–brand name and aspect accessibility should increase (and response latencies should decrease) when the new product is branded Big Red rather than when it is unbranded. When similarity is low, senior–brand name and aspect accessibility should decrease (and response latencies should increase) when the new product is branded Big Red rather than when it is unbranded. As we predicted, there is a significant category interaction between similarity and new product for all senior-brand accessibility measures (*p*s < .05).

Response latency means for the new-product planned contrasts (Big Red versus unbranded) appear in Table 1. Reinforcement is shown by a significant, positive contrast value, indicating that senior-brand associations are more accessible when the new product is a junior brand than when it is unbranded. Dilution is shown by a significant, negative contrast value, indicating that senior-brand associations are less accessible when the new product is a junior brand than when it is unbranded. As we predicted in H_{1a}, when category similarity is high, senior–brand name and aspect accessibility are reinforced; all four entries in Table 1, Column 1, are positive and significant (Fs > 5.00, *p*s < .05). When category similarity is low, these accessibilities are diluted; all entries in Table 1, Column 2, are negative and significant (Fs > 4.00, *p*s < .05), in support of H_{1b}.

TABLE 1
Study 1: Mean Change in Accessibility (Means)

	Category Similarity	
	Column 1: High (Bubble Gum)	Column 2: Low (Snack Bars)
Brand Name Accessibility (Aspect → Brand Name)		
Chewing gum → Big Red	212	–247
Cinnamon → Big Red	302	–276
Aspect Accessibility (Brand Name → Aspect)		
Big Red → chewing gum	327	–119
Big Red → cinnamon	174	–334

Notes. Mean changes in accessibility reflect our planned contrasts: response latency after exposure to unbranded product versus response latency after exposure to identical junior brand. Significant, positive and negative contrasts indicate reinforcement and dilution, respectively, for the senior brand. All contrasts are consistent with prediction and are significant at $p < .05$. a → b indicates that a is a retrieval cue for retrieving b.

Study 2: Effects of Category and Attribute Similarity on Response Latencies

The results of Study 1 confirm our predictions about the effects of category similarity. However, category similarity is only one of many ways that junior and senior brands may be similar or dissimilar. Consider again the example of Victor's Secret. Lingerie is a major part of its business; thus, many consumers might perceive category similarity between Victor's Secret and Victoria's Secret. At the same time, Victor's Secret may be viewed as different from Victoria's Secret on many attributes (e.g., "classiness," store appearance). What is the impact of this feature or attribute dissimilarity, and what happens when a junior brand operates in a dissimilar category but attempts to leverage distinctive attributes of the senior brand? For example, Sleep Inn was originally slated to be called McSleep Inn to communicate the "basic, convenient, inexpensive, and/or standardized qualities" associated with McDonald's (a dissimilar category with similar attributes; "McReading" 1988, p. 32). What is the impact of this feature or *attribute similarity*? Study 2 addresses these questions to capture the complexity of real-world situations.

Predictions

A variety of evidence suggests that category similarity is likely to take precedence over similarity in other aspects (e.g., attributes) in the determination of whether a junior brand dilutes a senior brand. Category (or relational) similarity, versus feature overlap (Tversky 1977), is likely to be accessed first and to have a greater influence on perceived similarity (Medin, Goldstone, and Gentner 1993). Thus, the encoding of new information (e.g., junior-brand knowledge) is likely to be guided by category similarity, such that closer associations are formed between senior and junior brands when category similarity is high. When category similarity is low, these associations should be more distinct and less tightly linked (Anderson, Green, and McCulloch 2000; Shiffrin 2003). Consistent with this view, it has been shown that category similarity influences the accessibility of constructs in memory more than does structural commonalities (attribute matching; Forbus, Gentner, and Law 1994). Therefore, our predictions are based on the assumption that category similarity takes precedence over attribute or feature similarity in the determination of how tightly linked junior-brand knowledge is with senior-brand knowledge in the shared brand network.

When category similarity is high and the senior and junior brands form a closely connected shared brand network, aspect similarity should have strong effects due to spreading activation. Similar aspects should reinforce senior–brand name and aspect accessibility, whereas dissimilar aspects should dilute brand name and aspect accessibility. Thus, senior-brand category associations (e.g., Big Red → chewing gum, chewing gum → Big Red) should be reinforced (because the categories are similar). However, senior-brand attribute associations (e.g., Big Red → cinnamon, and cinnamon → Big Red) should be reinforced when attribute similarity is high and diluted when attribute similarity is low.

When category similarity is low, senior-brand category associations should be diluted because the categories are dissimilar. Furthermore, because the senior and junior brands are encoded more distantly and distinctly in the shared brand network, similar attributes should have less ability to reinforce senior-brand attribute associations. Similar attributes are linked by their similarity, but because they are encoded in distinctive locations, the speed of activation of this link is lessened. Therefore, we expect that when attribute similarity is low, dilution of attribute associations is likely. When attribute similarity is high, reinforcement of attribute association is unlikely, though spreading activation along the relatively weak associations among similar attributes is likely to lessen dilution effects. (For all hypothesis tests, we measure both brand name and aspect accessibility; however, we expect no difference across these measures.)

H_2: Category and attribute similarity jointly determine the effect of a same-named junior brand on the senior brand's distinctive associations.

a. When category and attribute similarity are both high, the senior brand's category and attribute associations are reinforced (i.e., accessibilities increase).
b. When category similarity is high and attribute similarity is low, (i) the senior brand's category associations are reinforced (i.e., accessibilities increase), and (ii) the senior brand's attribute associations are diluted (i.e., accessibilities decrease).
c. When category similarity is low and attribute similarity is high, (i) the senior brand's category associations are diluted (i.e., accessibilities decrease), and (ii) the senior brand's attribute associations are not reinforced and may be diluted (i.e., accessibilities do not increase and may decrease).
d. When category and attribute similarity are both low, the senior brand's category and attribute associations are diluted (i.e., accessibilities decrease).

Method

Big Red chewing gum is again our senior brand. Our design is a 2 (category similarity: high [bubble gum] and low [snack bars]) × 2 (attribute similarity: high [cinnamon flavor and breath freshening] and low [strawberry flavor and containing immunity-boosting Echinacea]) × 2 (new product: junior brand [Big Red] and unbranded [Brand A]) × 2 (order of ad presentation) between-subjects design. Pretest participants indicated (on seven-point scales) that they are highly familiar with Big Red chewing gum (M = 6.73), strongly associate Big Red with fresh breath (M = 4.63) and cinnamon flavor (M = 6.52), and believe that it is unlikely that Big Red chewing gum is strawberry flavored (M = 1.41) or contains echinacea (M = 1.27). These likelihood ratings do not differ significantly from one another ($F_{1, 24}$ = 2.34, $p > .10$) and are both significantly lower than the likelihood that Big Red chewing gum freshens breath (M = 4.66) or is cinnamon flavored (M = 6.11; $ps < .01$). Participants were 235 undergraduate business students randomly assigned to conditions. Procedures and materials were otherwise identical to those of Study 1.

Results

Manipulation and data checks. As we expected, participants indicated (on seven-point scales) that they were highly familiar with Big Red chewing gum (M = 6.17) and strongly associated Big Red with fresh breath (M = 5.32) and cinnamon flavor (M = 6.21). Participants identified bubble gum and snack bars as high (M = 6.34) and low (M = 2.34), respectively, in similarity to chewing gum; this difference is significant ($F_{1, 212}$ = 11.66, $p < .05$). Participants also indicated that strawberry flavor (M = 1.73) and echinacea (M = 1.18) were equally weak in their association with the Big Red brand ($F_{1, 211}$ = .57, $p > .10$). A free-recall measure of ad content showed that participants were not confused about the relationship between the senior brand and the junior brand; only 1.3% of participants did not report the disclaimer, and this number did not vary by condition (χ^2_7 = 6.19, $p > .10$). Again, we collected a seven-point-scale measure of brand liking for the Big Red brand. Liking did not differ across conditions, indicating that our manipulations did not create tarnishment effects ($F_{7, 188}$ = 1.56, $p > .10$). The average error rate in the accessibility trials was low (2.3%) and did not vary by condition (χ^2_7 = 9.02, $p > .10$). As in Study 1, we eliminated participants who failed to recall the disclaimer and those who made errors.

Hypothesis tests. Again, we log-transformed response latencies to achieve distribution normality (Fazio 1990), and we used participants' mean response latency on filler trials as a covariate to isolate variance due to individual differences. Order of presentation of the filler advertisements had no impact on response latencies ($ps > .10$); thus, we do not discuss its effects further. Planned contrast values for the nontransformed response latencies appear in Table 2. Reinforcement (dilution) effects are indicated as significant and positive (negative) contrast values. All contrasts that are consistent with prediction appear in bold.

The results are as we predicted for 22 of the 24 contrasts. As H_{2a} predicts, when category and attribute similarity are both high, the senior brand is uniformly reinforced; all six entries in Table 2, Column 1, are positive and significant (Fs > 6.00, $ps < .05$). H_{2b} predicts that when category similarity is high and attribute similarity is low, category associations are reinforced and attribute associations are diluted. The results for category associations are mixed. As the first entry in Table 2, Column 2, shows, brand name accessibility for the category is reinforced, as predicted ($F_{1, 44}$ = 12.91, $p < .01$), but aspect accessibility for the category is not (Column 2, fourth entry; $F_{1, 42}$ = 2.07, $p > .10$). The results for attribute associations are all diluted; as we predicted, the second, third, fifth, and sixth entries in Table 2, Column 2, are negative and significant (Fs > 9.00, $ps < .01$).

$H_{2c(i)}$ predicts that when category similarity is low and attribute similarity is high, category associations are diluted. The results are mixed here as well. Brand name accessibility for the category is diluted (Table 2, Column 3, first entry; $F_{1, 44}$ = 12.91, $p < .01$), but aspect accessibility for the category is unaffected (Table 2, Column 3, fourth entry; $F_{1, 42}$ = 2.07, $p > .10$). $H_{2c(ii)}$ predicts that when cate-

TABLE 2
Study 2: Mean Change in Accessibility (Means)

	High Category Similarity (Bubble Gum)		Low Category Similarity (Snack Bars)	
	Column 1: High Attribute Similarity (Cinnamon, Fresh Breath)	**Column 2: Low Attribute Similarity (Strawberry, Echinacea)**	**Column 3: High Attribute Similarity (Cinnamon, Fresh Breath)**	**Column 4: Low Attribute Similarity (Strawberry, Echinacea)**
Brand Name Accessibility (Aspect → Brand Name)				
Chewing gum → Big Red	**195***	**243***	**–86***	**–234***
Cinnamon → Big Red	**152***	**–346***	**–144***	**–342***
Fresh breath → Big Red	**145***	**–66***	**95**	**–255***
Aspect Accessibility (Brand Name → Aspect)				
Big Red → chewing gum	**182***	–41	–132	**–402***
Big Red → cinnamon	**51***	**–327***	**–116***	**–316***
Big Red → fresh breath	**349***	**–250***	**63**	**–340***

*Contrast is significant at $p < .05$.

Notes: Mean changes in accessibility reflect our planned contrasts: response latency after exposure to unbranded product versus response latency after exposure to junior brand. Significant, positive and negative contrasts indicate reinforcement and dilution, respectively, for the senior brand. Bolded contrasts are consistent with prediction. a → b indicates that a is a retrieval cue for retrieving b.

gory similarity is low and attribute similarity is high, attribute associations are not reinforced and may be diluted. Consistent with this prediction, two of the four planned contrasts demonstrate dilution (Table 2, Column 3, second and fifth entries; Fs > 4.00, *ps* < .05), and two exhibit null effects (Table 2, Column 3, third and sixth entries; Fs < 2.20, *ps* > .05). Finally, as H_{2d} predicts, when category and attribute similarity are both low, the senior brand is uniformly diluted, as evidenced by the six negative and significant values in Table 2, Column 4 (Fs > 6.00, *ps* < .05).

Summary and Discussion of Studies 1 and 2

Dilution effects. Study 1 manipulated product category similarity and held attribute similarity constant at a moderate level. As we predicted, the accessibility of distinctive senior-brand associations was reinforced with high category similarity and was diluted with low category similarity. Study 2 manipulated both product category and attribute similarity. As predicted, in general, accessibility of the senior brand's category associations was reinforced when category similarity was high and was diluted when category similarity was low (though this effect did not hold for two aspect accessibility measures). As predicted, the accessibility of the senior brand's attribute associations depended on both category and attribute similarity. When category similarity was high, the senior brand's attribute associations were reinforced by attribute similarity and were diluted by attribute dissimilarity. When category similarity was low, the senior brand's attribute associations were not reinforced by attribute similarity; indeed, they were often diluted.

From a theoretical standpoint, our results suggest that category similarity determines the extent to which senior and junior brands are connected in the shared brand network. Further work examining the structure of the shared brand network can help better understand when junior brands pose the greatest threats and why. From managerial and public policy standpoints, our results offer insights into the understanding of present dilution and the future risk of dilution. Present dilution is likely when there are important dissimilarities between the senior and the junior brands, particularly at the category level. This insight is important because brands in different categories intuitively seem less of a risk, perhaps because they seem less likely to confuse consumers with respect to their (lack of) relationship with the senior brand. If confusion is indeed less likely, the dilution doctrine is an important source of trademark protection in this case because unlike infringement, it extends to situations in which there is no consumer confusion.

In contrast, the future risk of dilution is high when there are important similarities between the senior and the junior brands, particularly at the category level. Present dilution is unlikely—rather, the senior brand is likely to be reinforced—but the close association between the junior and the senior brands in the shared brand network means that future actions by the junior brand could create harmful associations that readily affect the senior brand (Swann and Davis 1994). In this situation, brand managers may find that proving infringement through evidence of confusion is more feasible. If a case for infringement cannot be made, careful monitoring for emerging evidence of dilution will be particularly important. It is noteworthy that similar junior brands (and risk of future dilution) may be common. As the shared brand network conceptualization predicts, in terms of the creation of desired associations, junior brands benefit the most when they are similar to the senior brand (Pullig, Simmons, and Netemeyer 2005).

Finally, it appears that the greatest challenges in detecting and proving present dilution occur when the junior brand is in a dissimilar category or possesses dissimilar attributes, but not both. In these cases, dilution is not uniform across associations, and some associations (e.g., the category association when category similarity it high) may be reinforced. Thus, dilution is more difficult to demonstrate. For the brand manager, careful measurement and an appreciation of which associations are likely rather than unlikely to be diluted are particularly important.

Response latency measure. The response latency measure is a good candidate for meeting the broad guidelines for evidence of dilution outlined by the Court. It goes beyond the idea of "mere association" between brands (rejected by the Court as inadequate) and shows an impact on how consumers process the senior brand, though it falls short of showing actual effects on sales (which the Court has deemed to be unnecessary). However, although response latencies provide a theoretically viable way to measure dilution, the procedures are cumbersome. Study 3 examines how well a simpler paper-and-pencil aided-recall task can replicate the effects of Study 2. Such a measure might provide a more tractable measurement methodology for the field.

Study 3: Effects of Attribute and Category Similarity on Aided Recall

We expect aided-recall measures to be less sensitive to differences in the accessibility of brand knowledge than response latency recognition measures. Aided recall requires the respondent to retrieve or recollect an association, whereas successful performance of a recognition task may be based solely on familiarity (Long and Prat 2002). Familiarity involves an assessment of the similarity between a target and the related knowledge in memory, not the explicit retrieval of the target from memory. Furthermore, familiarity may vary from weak to strong, whereas the retrieval required by recall is largely a threshold-based process. Therefore, although we expect that, in general, the aided-recall measure of dilution will replicate the effects we observed in Study 2, we also expect that it will show fewer of these effects.

Method

We use a paper-and-pencil aided-recall method to assess reinforcement and dilution across levels of category and attribute similarity. Our design is identical to that of Study 2, with one exception. In a within-subjects manipulation, we vary the order of recall: brand name recall first versus aspect recall first. We included this order manipulation for two reasons: First, because successful recall in a paper-and-pencil task is more cognitively and physically demanding, we were concerned about fatigue effects in reported recall. Second, because a recall task does not measure strength of association, just whether that strength exceeds the threshold necessary for recall, we were particularly concerned about differential transfer effects (e.g., giving participants Big Red as a retrieval cue for brand aspects increases the probability that they will subsequently recall Big Red when they are given an aspect as a retrieval cue).

Participants were 448 undergraduate business students randomly assigned to conditions. We asked them to evaluate at their own pace the same advertisements as those used in Study 2. Then, to provide closure to these processing instructions and to measure recall of the disclaimer, we asked participants to list all the aspects they remembered from the advertisements. Then, they participated in an unrelated study, which took approximately 15 minutes and included a series of arithmetic calculations on product search time and price expectations; simple arithmetic tasks are an effective way to clear short-term memory (Spears 1978). Finally, we administered the recall task.

We measured aspect accessibility by asking participants to recall all that came to mind when they thought of each brand from a list of brands, including the brand of interest (i.e., Big Red chewing gum). We measured brand name accessibility by asking them to recall what came to mind when they thought of a list of aspects, including the aspects of interest (i.e., chewing gum and cinnamon flavor). The order of aspect accessibility and of brand name accessibility measures was counterbalanced. Successful retrieval of the Big Red brand name in response to a distinctive aspect (e.g., chewing gum) or successful retrieval of a distinctive aspect (e.g., chewing gum) in response to the Big Red brand name is coded as a 1, and unsuccessful retrieval is coded as a 0. Two independent coders, who were unaware of the purpose of the research, coded the recall responses. Intercoder agreement was 94%. All differences were resolved through discussion. Following the recall task, participants completed manipulation and confounding check measures.

Results

Manipulation and data checks. Participants indicated (on seven-point scales) that they were highly familiar with Big Red chewing gum (M = 5.87) and strongly associated Big Red with cinnamon flavor (M = 6.35) and fresh breath (M = 4.91). They also identified bubble gum and snack bars as high (M = 6.07) and low (M = 2.84), respectively, in similarity to chewing gum; this difference is significant ($F_{1,480} = 43.67$, $p < .01$). Participants also rated strawberry flavor (M = 2.17) and echinacea (M = 2.22) as equally weak in their association with the Big Red brand ($F_{1,480} = 1.14$, $p > .10$). A free-recall measure of ad content showed that participants were not confused about the relationship between the senior brand and the junior brand; only 2.1% of participants did not report the disclaimer, and this number did not vary by condition ($\chi^2_{15} = 4.20$, $p > .10$). We eliminated participants who failed to recall the disclaimer. Order of presentation of the filler advertisements had no impact on proportions of successful recall (*p*s > .10); thus, we do not discus its effects further. Recall that task order resulted in significant differential transfer for all recall measures (*p*s < .05); therefore, the following analyses are based on the first recall task (brand name or aspect) for each respondent.

Hypothesis tests. As in Study 2, we test H_2 regarding the effects of category and attribute similarity on dilution and reinforcement of the senior brand. Our dependent measure is the successful recall of brand aspects, a nominal variable; thus, we use a logistic regression model. As in Study 2, our planned contrasts compare participants who were exposed to the junior brand with those who were exposed to an identical unbranded product. Contrast values and significance tests appear in Table 3. Contrasts that are consistent with the predictions in H_2 appear in bold. In Study 2, using response latency measures, we found that 22 of the 24 contrasts were consistent with prediction. In Study 3, using an aided-recall measure, we found that though all contrasts were in the predicted direction, only 16 of the 24 contrasts were consistent with prediction when statistical significance is considered.

H_{2a} predicts that when category and attribute similarity are both high, the senior brand is reinforced. As we show in Table 3, Column 1, four of the six accessibility measures are consistent with this prediction (Wald statistics > 3.41, $p < .05$). H_{2b} predicts that when category similarity is high and attribute similarity is low, the senior brand's category associations are reinforced. This prediction holds for only one of the two measures; when category is used as a retrieval cue, brand name accessibility is reinforced (chewing gum → Big Red; Wald statistic = 3.98, $p < .05$). H_{2b} also predicts that when category similarity is high and attribute similarity is low, attribute associations are diluted. This prediction holds for only one of the four measures. When the brand name is used as a retrieval cue, aspect accessibility for cinnamon flavor is diluted (Big Red → cinnamon flavor; Wald statistic = 4.07, $p < .05$).

$H_{2c(i)}$ predicts that when category similarity is low and attribute similarity is high, category associations are diluted. This prediction is supported (Column 3, first and fourth entries; Wald statistics > 3.71, $p < .05$). $H_{2c(ii)}$ predicts that attribute associations are not reinforced and may be diluted in this situation. Consistent with prediction, aspect accessibility for cinnamon flavor is diluted (Big Red → cinnamon flavor; Wald statistic = 6.74, $p < .01$), and all other attribute associations in this condition are unaffected. Finally, H_{2d} predicts that when both category and attribute similarity are low, the senior brand is diluted. Table 3, Column 4, supports this prediction for each of the aspect accessibility measures (fourth, fifth, and sixth entries; Wald statistics > 4.00, $p < .05$) and for brand name accessibility when the category is used as a retrieval cue (chewing gum → Big Red; Wald statistic = 4.07, $p < .05$).

Discussion

As we expected, recall measures of dilution were less sensitive to differences in accessibility than the response latency recognition measure. However, all contrasts in Table 3 are in the predicted direction, and 67% meet our statistical criteria (recall that some null effects are predicted); this recall measure is highly correlated with our response latency measure from Study 2 (r = .75, $p < .001$). Thus, the data are encouraging about the viability of a recall measure. However, these results emphasize the need to consider multiple measures in the assessment of dilution. For example, it may make sense to trade off sensitivity for convenience and use a recall measure for general monitoring purposes but response latencies for a more in-depth understanding or when proof of dilution is crucial.

Study 4: Effects of a Junior Brand on Consideration and Choice Probabilities

From a practical perspective, dilution is of interest because it can affect sales in the marketplace. Thus, a valid measure of dilution captures changes in brand knowledge that lead to reduced consideration and choice of the senior brand. Theoretically, the reduced brand name and aspect accessibility we observed using response latency and recall measures have such effects. We now examine whether situations that lead to dilution, as indicated by response latency and recall measures, reduce the probability of consideration and

TABLE 3
Study 3: Change in Proportion Recalling Brand Name/Aspects

	High Category Similarity (Bubble Gum)		Low Category Similarity (Snack Bars)	
	Column 1: High Attribute Similarity (Cinnamon, Fresh Breath)	Column 2: Low Attribute Similarity (Strawberry, Echinacea)	Column 3: High Attribute Similarity (Cinnamon, Fresh Breath)	Column 4: Low Attribute Similarity (Strawberry, Echinacea)
Brand Name Accessibility (Aspect → Brand Name)				
Chewing gum → Big Red	**.30***	**.29***	**–.21***	**–.23***
Cinnamon → Big Red	.12	.01	**–.06**	–.10
Fresh breath → Big Red	.05	–.02	**–.04**	–.05
Aspect Accessibility (Brand Name → Aspect)				
Big Red → chewing gum	**.23***	.19	**–.20***	**–.23***
Big Red → cinnamon	**.19***	**–.27***	**–.31***	**–.36***
Big Red → fresh breath	**.22***	–.01	**–.03**	**–.12***

*Contrast is significant at $p < .05$.

Notes. Contrast values reflect the proportion successfully recalling brand name/aspects after exposure to junior brand versus the proportion successfully recalling brand name/aspects after exposure to control brand. Positive (negative) values indicate reinforcement (dilution) effects. Bolded contrasts are consistent with prediction. a → b indicates that a is a retrieval cue for retrieving b.

choice. We focus on the situation that has most consistently shown evidence of dilution through accessibility measures, that is, when both attribute and category similarity are low.

H_3: When category similarity and attribute similarity are low, given distinctive aspects as a choice criterion, the probability of the senior brand being included in a consideration set decreases.

H_4: When category similarity and attribute similarity are low, given distinctive aspects as a choice criterion, the probability of the senior brand being chosen decreases.

Method

The design is a 3 (senior brands: Big Red, Gap, and Trix) × 2 (condition: diluting versus control) factorial design in which we manipulate the senior brand between subjects and the condition within subjects. Participants were 219 undergraduate students who participated in one diluting condition and two control conditions. As in our previous studies, we achieved manipulations through ad exposures using a computer-administrated procedure. Advertisements in the diluting condition were for a junior brand in a dissimilar category and with dissimilar attributes (e.g., when Big Red was the senior brand, the junior brand was Big Red strawberry snack bars). Advertisements in the control conditions were identical except that we identified the brand as a well-known brand in the category (e.g., when Big Red was the senior brand, the advertisement in the control condition was for Nutri-Grain strawberry snack bars). After participants were exposed to the advertisements, we asked them to choose a brand in each senior-brand category, given distinctive aspects of the senior brand as a choice criterion. We examine how consideration and choice of the senior brand are affected by exposure to a junior brand (versus exposure to an identical product that does not use the senior brand's name).

Previous results showed that students are highly familiar with Big Red chewing gum, that cinnamon flavor is strongly associated with Big Red, and that the snack bar category and strawberry flavor are viewed as dissimilar to Big Red's distinctive aspects. Therefore, one of our junior brands is Big Red strawberry snack bars, and our choice criterion is long-lasting cinnamon flavor. Furthermore, pretest participants indicated (on seven-point scales) a high level of familiarity with Gap clothing (M = 6.82) and Trix cereal (M = 6.22); that cotton and khakis are strongly associated with Gap (M = 6.10 and 6.52, respectively); and that colorful, fun, fruity, and for kids are strongly associated with Trix (M = 6.40, 5.20, 6.06, and 6.56, respectively). Pretest participants further indicated that the bed sheets category is dissimilar to clothing (M = 2.04) and that the cake mix category is dissimilar to cereal (M = 2.21). Finally, pretest participants indicated that stain resistant is an unlikely attribute of Gap clothing (M = 3.13) and that low fat is an unlikely attribute of Trix cereal (M = 2.13). Therefore, our two other junior brands are Trix low-fat cake mix and Gap stain-resistant sheets. Our clothing choice criterion is "a good source for cotton khakis," and our cereal choice criterion is "a colorful, fun, fruity cereal for kids."

After viewing the advertisements, participants engaged in practice trials to learn the computer-administered choice task and then took part in the main experiment. The choice task was a simulated shopping experience. The brands were represented in color as brand logos or packages of equal size. Participants were asked to choose a set of up to four brands that they would consider, given the choice criterion. Then, this set of brands was displayed on the screen, and participants were asked to choose the brand option that best satisfied the choice criterion. In each step of the simulated choice process, the brands were displayed on the screen in random order. This was done for each of the three senior-brand categories. Following the choice task, participants were asked to recall ad content and to respond to manipulation and confounding check measures. As in the prior studies, we deleted any respondent who failed to report the disclaimer (1.2%); this did not vary by condition ($\chi^2_2 = 1.52$, $p > .10$).

Results and Discussion

The results appear in the top half of Table 4. When exposure to the junior brand resulted in significantly reduced probability of consideration or choice, significant differences in proportions appear in bold. In support of H_3, the likelihood

TABLE 4

Studies 4 and 5: Proportion of Respondents Considering and Choosing Senior Brand

	Consideration		Choice	
	Exposed to Junior Brand?		Exposed to Junior Brand?	
Senior Brand	No	Yes	No	Yes
Study 4: No Delay				
Big Red	**.85**	**.52**	**.60**	**.15**
Trix	.72	.55	**.31**	**.08**
Gap	**.68**	**.49**	**.25**	**.09**
Study 5: Five-Day Delay				
Big Red	.60	.48	.43	**.22**
Trix	**.69**	**.52**	.33	.31
Gap	**.45**	**.29**	.38	**.17**

Notes: Bolded pairs of proportions within the "Considered" and "Choice" columns are significantly different at $p <$ 05 and are consistent with prediction.

of the senior brand being included in a consideration set was significantly decreased with the introduction of a dissimilar junior brand for two of the three brands ($\chi_1^2 = 17.56$, $p < .01$, for Big Red; and $\chi_1^2 = 4.55$, $p < .05$, for Gap). In support of H_4, the probability of choosing the senior brand was significantly reduced in all three categories ($\chi_1^2 = 19.87$, $p < .01$, for Big Red; $\chi_1^2 = 4.18$, $p < .05$, for Trix; and $\chi_1^2 = 3.89$, $p < .05$, for Gap). Thus, conditions that resulted in measured dilution when we used accessibility measures also resulted in decreased consideration and choice of the senior brand, as we predicted. Furthermore, these results have generality across brands.

Study 5: Consideration and Choice Effects with Delay

In each of the four preceding studies, we examined dilution effects shortly after exposure to the junior brand. In Study 5, we examined whether effects on consideration and choice probabilities occur when there is a substantial delay between exposure to the junior brand and the choice involving the senior brand. The design and procedures are identical to those of Study 4, except that the procedure is administered with paper and pencil, and there is a five-day delay between exposure to the diluting advertisements and the choice task. Participants were 49 randomly assigned undergraduate students.

The results appear in the bottom half of Table 4. Significant differences appear in bold. Consideration and choice effects are apparent even with a five-day delay between exposure to the junior brand and the choice task. Exposure to the junior brand significantly decreased consideration probability for two of the three brands ($\chi_1^2 = 4.93$, $p < .05$, for Trix; and $\chi_1^2 = 3.63$, $p < .05$, for Gap) and also decreased choice probability for two of the three brands ($\chi_1^2 = 5.39$, $p < .05$, for Big Red; and $\chi_1^2 = 4.84$, $p < .05$, for Gap). Together, Studies 4 and 5 provide strong evidence of the power of a junior brand to decrease the consideration and choice of a senior brand.

General Discussion

We focused on two questions of interest to brand managers and public policy makers: First, how can and should trademark dilution be measured? Second, when is it likely to occur? We draw on our results to address these questions and identify important research issues.

How Can and Should Trademark Dilution Be Measured?

We defined blurring dilution as a weakening of the associations between a brand and its distinctive aspects (e.g., category, attributes, benefits) that define the meaning of the brand in consumers' minds and, at a more concrete level, as a reduction in brand name and aspect accessibility. We then operationalized dilution as response latency to recognize distinctive aspects (Studies 1 and 2) and as aided recall of distinctive aspects (Study 3). In Studies 4 and 5, we did not directly measure dilution but rather its predicted effects on consideration and choice.

Construct validity. This work provides strong evidence of the construct validity of our measures. First, our measures have content, or face, validity. Definitions of blurring center on the loss of brand meaning coupled with the probability of reduced economic value, particularly through lost sales (Federal Trademark Dilution Act 1995; Schechter [1927] 1970; Simonson 1993). Our accessibility measures capture losses in brand meaning that are theoretically linked to reduced brand choice. Less accessible brand names are less likely to be recalled and less likely to enter a consumer's consideration set when they are encountered in the shopping environment. Less accessible aspects are less confidently associated with the brand in choice situations (Alba, Hutchinson, and Lynch 1991).

Second, we provide evidence of convergent and discriminant validity. Our response latency recognition measure is highly correlated with our aided-recall measure, but it is uncorrelated with brand liking, ensuring that we are measuring blurring dilution rather than tarnishment.

Finally, our measures have nomological validity (Cronbach and Meehl 1955); that is, the measured construct behaves as it is predicted to behave. As we predicted, category and attribute similarity moderate measured dilution. Furthermore, the low-similarity situations that lead to measured dilution also lead to decreased consideration and choice.

Viability as legal evidence. Our measures also fall within the guidelines set by the Court; they go beyond mere association but do not extend to actual lost sales, and thus they are potentially useful for brands seeking protection from dilution (Morrin and Jacoby 2000). Response latencies are likely the most sensitive measure of dilution, but accessibilities can also be measured in ways that are more amenable to field survey methods (e.g., aided recall; Peterson, Smith, and Zerrillo 1999). Furthermore, as more is learned about individual differences in dilution effects—for example, people with higher brand familiarity and category knowledge (Morrin and Jacoby 2000) may exhibit smaller effects—this knowledge may be taken into account when selecting an appropriate dilution measure.

Blurring versus tarnishment. Although we isolated blurring dilution in these experimental studies, blurring and tarnishment may occur together. Junior-brand associations may be both unattractive and dissimilar to those of the senior brand. Furthermore, severe blurring may itself lower brand evaluations. Therefore, as a practical matter, firms may wish to assess multiple types of negative junior-brand effects routinely: blurring of distinctive aspects, accretion of unattractive associations, and lowered brand evaluations. The kinds of accessibility measures we examined herein should prove particularly useful for assessing the extent to which negative junior-brand aspects have become associated with the senior brand (Morrin and Jacoby 2000).

When Is Dilution Likely?

Our results suggest that blurring dilution is least likely in copycat strategies, that is, when a junior brand emerges in a similar category to the senior brand and with other similar aspects (e.g., high attribute similarity). This finding contradicts what is often brand management's gut reaction: that highly similar junior brands are the most serious threats. However, we agree that similar junior brands create an exposure for the senior brand. By virtue of its strong interconnections with the senior brand in the shared brand network, the junior brand has the power to dilute a senior brand through any future actions that create dissimilar associations (Pullig, Simmons, and Netemeyer 2005; Swann and Davis 1994). In contrast, our data show that junior brands in similar categories but with dissimilar associations can dilute senior-brand knowledge.

Case law implicitly recognizes the danger from junior brands in similar categories (Vapnyar 2003). However, the legal requirement for evidence of dilution means that the senior brand may be unable to protect itself preemptively from the risk of future dilution by a highly similar junior brand. Fortunately, measurable dilution is likely to be found in all other situations we examined. Our results provide some information about where to expect dilution (e.g., not for category associations in similar categories) and where effect sizes are likely to be smallest (e.g., dissimilar category, similar attributes) to guide the measurement process.

Further Research

Three issues deserve special attention: First, how similar must trademarks be to create dilution? The Ninth Circuit recently ruled that Orbit bicycles were not diluted by the junior brand OrbiTrek exercise bicycles, because the marks were not sufficiently similar (*Thane Int'l Inc. v. Trek Bicycle Corp.* 2002). Brand managers should be concerned with this line of reasoning. Judgments about what constitutes sufficient similarity should be guided by an in-depth understanding of the dimensions of similarity (e.g., visual, conceptual, lexical, phonological) that link a junior brand with a senior brand. Such knowledge could also alert brand managers to junior brands that offer greater or lesser dilution exposure to the senior brand.

Second, are some types of trademarks more susceptible to dilution than others? For example, a recent trend is to create "fanciful" or "coined" brand names (e.g., Verizon, Accenture), which do not have suggestive associations with the product they mark. An advantage of fanciful marks is that their associations are (at least initially) under the control of the firm, but are they more or less susceptible to dilution than suggestive marks, such as Drano or Sonicare, which convey product associations? Recent work on memory suggests that the representation of these two types of marks differs such that we would expect fanciful names to be more susceptible to dilution of brand name accessibility than suggestive names (for a similar idea in a somewhat different domain, see Young 2001, cited in Schacter 2001). Furthermore, the language-based associations of suggestive names may reduce overall susceptibility to dilution. If these predictions prove correct, they have three implications for defending against dilution: First, a brand's susceptibility may be reduced by designing trademarks that facilitate communication of the brand's distinctive aspects. Second, for fanciful marks, resources may be allocated more efficiently by focusing on increasing brand name accessibility, which is more often diluted. Third, it is possible to monitor for dilution more effectively by using multiple measures. Brand name accessibility dilution may occur even when there is no aspect accessibility dilution.

Third, to predict dilution effects correctly, brand managers and public policy makers alike require an understanding of the structure of the shared brand network. On the basis of recent memory research, we assumed that junior brands are encoded proximate to and well integrated with senior-brand knowledge when category similarity is high and more distinctly and distantly from senior-brand knowledge when category similarity is low. In general, the accessibility effects we observed were consistent with this model. However, we did not directly access memory structure. Furthermore, there are many variables that seem likely to alter the degree of interconnectedness of the shared brand network. For example, attribute dissimilarity can manifest in ways that may differ in their capacity to create the close linkages that lead to dilution. Some dissimilar attributes are different points along the same dimension (i.e., cinnamon and strawberry flavors), and others reflect completely different dimensions (i.e., cinnamon flavor and vitamin enriched). It might be expected that the former creates a closer linkage and, thus, greater dilution.

Back to Victor's/Victoria's Secret

Our results suggest that Victoria's Secret may not be much harmed in the short run by Victor's Secret, especially if most consumers who are aware of Victor's Secret know that it sells only "adult" lingerie. This category association is similar to that of Victoria's Secret (i.e., sexually alluring lingerie), though there may be a clear distinction on the "tastefulness" or "product offering" dimensions. Thus, Victoria's Secret may wish to focus on tarnishment or dilution of attribute-level aspects in its attempt to demonstrate harm.

Perhaps a greater concern is the risk of future dilution from this junior brand. Whereas a single "adult" shop named Victor's Secret would seem insignificant in the face of Victoria's Secret's marketing power—simply a joke to be appreciated—the risk is real. One of the values of building equity around an arbitrary mark such as Victoria's Secret (i.e., one that consists of common words that do not suggest or describe the product) is that these marks are the easiest to protect legally from infringement (Launer 1996). If others can use the mark or a parody version of it, the value of building an arbitrary brand is lessened. Although current trademark dilution law does not offer legal remedies for the risk of future dilution, senior brands with nondiluting junior brands would be well advised to monitor the activities of these unintended associates.

REFERENCES

Alba, Joseph W., J. Wesley Hutchinson, and John G. Lynch Jr. (1991), "Memory and Decision Making," in *Handbook of Consumer Behavior*, Thomas S. Robertson and Harold H. Kassarjian, eds. Englewood Cliffs, NJ: Prentice Hall, 1–49.

American Express Co. v. Vibra Approved Lab Corp. (1989), 10 U.S.P.Q. 2d (BNA) 2006, 2013-14 (S.D.N.Y.).

Anderson, John R. (1983), "A Spreading Activation Theory of Memory," *Journal of Verbal Learning and Verbal Behavior*, 22 (3), 261–95.

Anderson, Michael C., Collin Green, and Kathleen C. McCulloch (2000), "Similarity and Inhibition in Long-Term Memory: Evidence for a Two-Factor Theory," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26 (5), 1141–59.

Burke, Raymond R. and Thomas K. Srull (1988), "Competitive Interference and Consumer Memory for Advertising," *Journal of Consumer Research*, 15 (June), 55–68.

Cronbach, Lee J. and Paul E. Meehl (1955), "Construct Validity in Psychological Tests," *Psychological Bulletin*, 52 (July), 281–302.

Fazio, Russell H. (1990), "A Practical Guide to the Use of Response Latency in Social Psychological Research," in *Review of Personality and Social Psychology*, Vol. 11, Clyde Hendrick and Margaret S. Clark, eds. Beverly Hills, CA: Sage Publications, 74–97.

Federal Trademark Dilution Act (1995), 15 U.S.C. 1125 (c) (1), § 43 (c) (1).

Feldwick, Paul (1996), "What Is Brand Equity Anyway, and How Do You Measure It?" *Journal of the Market Research Society*, 38 (2), 85–104.

Forbus, Kenneth D., Dedre Gentner, and Keith Law (1994), "MAC/FAC: A Model of Similarity-Based Retrieval," *Cognitive Science*, 19 (2), 141–205.

Humphreys, Michael S., Gerald Tehan, Annissa O'Shea, and Scott W. Bolland (2000), "Target Similarity Effects: Support for the Parallel Distributed Processing Assumptions," *Memory and Cognition*, 28 (July), 798–821.

Jacoby, Jacob (2001), "The Psychological Foundations of Trademark Law: Secondary Meaning, Genericism, Fame, Confusion, and Dilution," *The Trademark Reporter*, 91 (5), 1013–1071.

Johar, Gita Venkataramani and Carolyn J. Simmons (2000), "The Use of Concurrent Disclosures to Correct Invalid Inferences," *Journal of Consumer Research*, 26 (March), 306–321.

John, Deborah Roedder, Barbara Loken, and Christopher Joiner (1998), "The Negative Impact of Extensions: Can Flagship Products Be Diluted?" *Journal of Marketing*, 62 (January), 19–32.

Keller, Kevin Lane (1993), "Conceptualizing, Measuring, and Managing Customer-Based Brand Equity," *Journal of Marketing*, 57 (January), 1–22.

Launer, Christine (1996), "More Than Name at Stake When Choosing Trademark," *Puget Sound Business Journal*, (June 28), (accessed November 21, 2005), [available at http://www.bizjournals.com/seattle/stories/1996/07/01/smallb9.html].

Lee, Thomas R. (2004), "Demystifying Dilution," *Boston University Law Review*, 84 (October), 1–94.

Long, Debra L. and Chantel S. Prat (2002), "Memory for *Star Trek*: The Role of Prior Knowledge in Recognition Revisited," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28 (November), 1073–2082.

McCarthy, J. Thomas (2004), "Dilution of a Trademark: European and United States Law Compared," *The Trademark Reporter*, 94 (6), 1163–81.

"McReading" (1988), *Harper's Magazine*, (December), 32.

Mead Data Central Inc. v. Toyota Motor Sales, USA Inc. (1989), 875 F. 2d 1026 (2d Cir.).

Medin, Douglas L., Robert L. Goldstone, and Dedre Gentner (1993), "Respects for Similarity," *Psychological Review*, 100 (2), 254–78.

Morrin, Maureen (1999), "The Impact of Brand Extensions on Parent Brand Memory Structures and Retrieval Processes," *Journal of Marketing Research*, 36 (November), 517–25.

——— and Jacob Jacoby (2000), "Trademark Dilution: Empirical Measures for an Elusive Concept," *Journal of Public Policy & Marketing*, 19 (Fall), 265–76.

Mosely v. V. Secret Catalogue Inc. (2003), 537 U.S. 418.

Nike v. Variety Wholesalers Inc. (2003), 274 F.Supp. 2d 1352 (S.D. Ga.).

Peterson, Robert A., Karen H. Smith, and Philip C. Zerrillo (1999), "Trademark Dilution and the Practice of Marketing," *Journal of the Academy of Marketing Science*, 27 (2), 255–68.

Pullig, Chris, Carolyn J. Simmons, and Richard G. Netemeyer (2005), "Free-Riding Junior Brands: Effects and Policy Issues," working paper, Hankamer School of Business, Department of Marketing, Baylor University.

Savin Corp. v. The Savin Group (2003), WL 22451731, 10 (S.D.N.Y.).

Schacter, Daniel L. (2001), *The Seven Sins of Memory: How the Mind Forgets and Remembers*. Boston: Houghton Mifflin.

Schechter, Frank I. ([1927] 1970), "The Rational Basis of Trademark Protection," *The Trademark Reporter*, 60 (3), 334–52. (Originally published in *Harvard Law Review*, 40 (6), 813–33.)

Shiffrin, Richard M. (2003), "Modeling Memory and Perception," *Cognitive Science*, 27 (3), 341–78.

Simonson, Alexander (1993), "How and When Do Trademarks Dilute? A Behavioral Framework to Judge 'Likelihood' of Dilution," *The Trademark Reporter*, 83 (2), 149–74.

Spears, Norman E. (1978), *The Processing of Memories: Forgetting and Retention*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Swann, Jerre B. and Theodore H. Davis Jr. (1994), "Dilution an Idea Whose Time Has Gone: Brand Equity as Protectable Property, the New/Old Paradigm," *The Trademark Reporter*, 84 (3), 267–99.

Thane Int'l Inc. v. Trek Bicycle Corp. (2002), 305 F. 3d 894 (9th Cir.).

Tversky, Amos (1977), "Features of Similarity," *Psychological Review*, 84 (4), 327–52.

Vapnyar, Vadim (2003), "Moseley v. V Secret Catalogue Inc.: Redefining the Scope of the Federal Trademark Dilution Act," *St. John's Law Review*, 77 (Summer), 675–96.

Xueming Luo, Rebecca J. Slotegraaf, & Xing Pan

Cross-Functional "Coopetition": The Simultaneous Role of Cooperation and Competition Within Firms

Extant marketing literature tends to view cross-functional relationships as primarily cooperative or competitive in nature, but not both. In contrast, this research focuses on cross-functional "coopetition" (i.e., the joint occurrence of cooperation and competition across functional areas within a firm). Using responses from midlevel managers and top executives, the authors find that cross-functional coopetition enhances a firm's customer and financial performance. The authors further show that this influence is mediated by market learning, indicating that performance returns to cross-functional coopetition occurs through an underlying learning mechanism.

Scholars in marketing and strategic management argue that knowledge transfer across functional boundaries is critical for numerous outcomes, including new product success (e.g., Griffin and Hauser 1992), organizational learning (Huber 1991), and overall firm performance (e.g., Gray and Meister 2004). In the marketing literature, Menon and Varadarajan (1992) argue that a firm's market knowledge must be transferred or disseminated across departments before knowledge can play a critical role as a strategic asset.[1] Research also indicates that a firm's competitive advantage lies in its ability to transfer market knowledge across departments (e.g., Maltz and Kohli 1996).

However, the transfer of knowledge across functional boundaries can be rather difficult and complicated. Even the most organized efforts to share knowledge are often impeded by employees' tendencies to guard and selectively share information (Gilmour 2003). Consider the number of companies that have experienced difficulties with internal knowledge transfer. For example, Hewlett-Packard's early failure in the laptop market with a 23-pound product can be attributed to its lack of knowledge flow between its marketing and engineering departments (Fisher, Maltz, and Jaworski 1997). General Motors found that its knowledge-sharing attempts in the transfer of manufacturing insights from its Saturn division to its other divisions proved more difficult than expected (Kerwin and Woodruff 1992). Barilla SpA's own sales and marketing areas impeded knowledge transfer to its operations area when it attempted to install a just-in-time distribution system (Hammond 1994).

Within academic research across marketing and management, scholars also recognize that knowledge transfer is valuable but often difficult to achieve (e.g., Kogut and Zander 1992; Szulanski 1996) and have subsequently studied barriers to knowledge transfer. Although research points to several factors that can impede knowledge transfer across functional boundaries within an organization, a critical factor is the degree of competition across functions. In particular, research shows that an arduous relationship between the source and the recipient (Szulanski 1996), or interfunctional rivalry (Maltz and Kohli 1996), can impede internal knowledge transfer. In numerous situations, functions must compete for a firm's scarce resources and thus are often reluctant not only to share information to prevent a competing function from gaining knowledge but also to receive information for fear that it will heighten the value of a competing function's knowledge (Gupta and Govindarajan 2000). Thus, the question is whether competing departments can effectively cooperate with one another to enhance organizational learning and performance. More generally, how should firms strategically manage cross-functional competition and cooperation to achieve competitive advantage?

The extant marketing literature has not addressed this important question, because it is often assumed that cross-functional interactions between marketing and other functional units are primarily cooperative or competitive, but not both. On the one hand, many researchers have viewed cross-functional interactions from a cooperation aspect, focusing on the value of interfunctional coordination and communication (e.g., Griffin and Hauser 1992; Moorman and Rust 1999; Narver and Slater 1990). On the other hand, others have viewed cross-functional interactions from a competition aspect, focusing on interdepartmental rivalry

---

[1] We refer to "functions" and "departments" interchangeably throughout this article.

Xueming Luo is Assistant Professor of Marketing, College of Business Administration, University of Texas, Arlington (e-mail: luoxm@uta.edu). Rebecca J. Slotegraaf is Assistant Professor of Marketing and Lilly Faculty Fellow (e-mail: rslotegr@indiana.edu), and Xing Pan is Assistant Professor of Marketing (e-mail: xpan@indiana.edu), Kelley School of Business, Indiana University. The authors appreciate the constructive comments and suggestions from the anonymous *JM* reviewers, Rohit Deshpandé, Ken Price, K. Sivakumar, and seminar participants at the University of Texas, Arlington, and Jinan University in China on previous versions of this manuscript.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

(e.g., Homburg, Workman, and Krohmer 1999; Levitt 1969). However, although functional areas compete with one another in the pursuit of divergent individual goals and strategic priorities (e.g., Houston et al. 2001), they must also cooperate so that they work toward the firm's common interests (Narver and Slater 1990). Consequently, an organization's functional areas are often forced to compete and cooperate simultaneously with one another. Thus, it is critical to understand how these seemingly conflicting relationships interplay and affect firm performance.

In an effort to understand these effects better, we focus on cross-functional "coopetition," which we define as the joint occurrence of cooperation and competition across functional areas within a firm. A coopetition view of the interactions between marketing and other functional areas not only emphasizes the need to address complex interdepartmental structures but also echoes the importance of coordinating diverse and competitive knowledge and skills across the cross-functional boundaries within a firm. We argue that cross-functional coopetition improves a firm's customer and financial performance.

We also investigate whether market learning plays a mediating role in the relationship between cross-functional coopetition and firm performance. This mediating role would suggest that there is an underlying learning mechanism by which cross-functional coopetition contributes to superior performance. Market learning is grounded in the knowledge-based view of the firm (Grant 1996) and is recognized as an important source for a firm's competitive advantage (e.g., Kogut and Zander 1992; Nonaka 1994; Sinkula 1994). Although marketing scholars have focused on the value of knowledge across a broad array of topics, including the utilization of market knowledge (e.g., Deshpandé and Zaltman 1982; Menon and Varadarajan 1992; Moorman, Zaltman, and Deshpandé 1992) and the exchange of knowledge across functional units (Fisher, Maltz, and Jaworski 1997; Griffin and Hauser 1992; Maltz and Kohli 1996), how intrafirm cooperation among competing functions influences a firm's learning about key marketing activities has yet to be explored. Following prior studies (e.g., Menon et al. 1999; Moorman 1995; Srivastava, Shervani, and Fahey 1998), we refer to market learning as an organization's knowledge stores of key marketing activities, such as developing new products, building brand image among customers, and establishing channel partnerships and networks in the market.

Therefore, in this study, we empirically test (1) whether the joint occurrence of cross-functional competition and cross-functional cooperation (intensity and ability) improves a firm's customer and financial performance and (2) whether a firm's market learning mediates these performance returns. Our results show that cross-functional coopetition has an important effect on performance outcomes through enhanced market learning, paving the way for new insight into how cross-functional interactions can affect a firm's competitive advantage. Our study also offers managers evidence that cooperation and competition should *both* be strategically stimulated across functions to promote intrafirm knowledge transfer and to enhance the firm's customer and financial performance.

Cross-Functional Coopetition

Theoretical Nature of Cross-Functional Coopetition

Coopetition, or the joint occurrence of cooperative and competitive behaviors, can exist at multiple levels, including firms, strategic business units (SBUs), departments, and task groups (e.g., Brandenburger and Nalebuff 1996; Hamel, Doz, and Prahalad 1989; Tsai 2002). One theoretical foundation of coopetition can be drawn from research on social structure in the sociology literature. In particular, the social embeddedness framework purports that relations are always present and that the social structure of these relations influences subsequent behaviors (Granovetter 1985; Uzzi 1997, 1999). In consideration of the types of social relations, weak ties (Granovetter 1973) are characterized by sporadic interactions, yet they can offer high returns by linking people or firms to diverse pools of information (Burt 1992). In contrast, embedded ties are characterized by frequent and stronger interactions such that information is perceived as more trustworthy (Granovetter 1985) and cooperation is high (Gulati 1998). In considering both types of relations, research shows that the greatest value is recognized when there is a complementary mix of both forms (Uzzi 1999). Thus, we expect that firms enhance their performance when they exhibit cooperative social ties that are nested in a broader competitive framework.

Research in the marketing literature supports and extends this notion. For example, Vargo and Lusch (2004) point to the importance of collaboration not only among interorganizational partners but also among intraorganizational functions for a firm's long-term viability. At the intraorganizational level, some marketing scholars have also recognized that interdepartmental interaction may be a double-edged sword, involving both collaboration and rivalry (e.g., Ruekert and Walker 1987). We explicitly address this interdepartmental interaction by examining the effects of cross-functional coopetition, or the joint occurrence of cooperation and competition, across functional areas within a firm.

In cross-functional coopetition, interdepartmental interactions and the transfer of knowledge across functional areas can be both competitive and cooperative in nature. The competitive nature often occurs because knowledge can generate private gains for individual departments to outperform their counterparts. Indeed, cross-functional competition may result from direct comparisons among functional units (Levitt 1969; Maltz and Kohli 1996), interdepartmental struggles to obtain limited tangible resources (e.g., organizational capital, personnel) and intangible resources (e.g., top executives' mental time, attention) (Frankwick et al. 1994), and divergent goals and strategic priorities (Houston et al. 2001; Ruekert and Walker 1987). At the same time, knowledge transfer across functions within a firm may be cooperative in nature because departments need to collaborate with one another in the transfer of market knowledge for the common interests of the firm. Informal cooperative interaction among functional departments is particularly important for obtaining more frequent and valuable information (Granovetter 1985) and for achieving greater knowledge transfer across departments (Ruekert and Walker 1987). For example, "hall talk" and storytelling are informal knowledge transfer processes that occur among departments as a means to bring people and departments together and to facilitate attainment of the organization's goals (Kohli and Jaworski 1990). Thus, the degree to which a firm's departments cooperate in conjunction with various levels of competition in the firm's social structure jointly defines the firm's level of cross-functional coopetition.

To clarify the concept of cross-functional coopetition further, we briefly discuss how it differs from two related constructs: First, cross-functional coopetition is similar to interorganizational alliances in that both involve cooperation between competing segments (i.e., functions, firms), yet important differences exist. An obvious difference is the level of analysis; the former is focused at the intraorganizational level, and the latter is focused at the interorganizational level. As a result, for example, the competitive element of coopetition within a firm is likely to be different from the competitive element between firms, such that rivalry and conflict are likely to be less extreme within a firm than between firms. In addition, underlying characteristics of interorganizational alliances, such as anticipated future interaction (Heide and Miner 1992) and relationship age (Anderson and Weitz 1989), are not likely to play a role in cross-functional coopetition. Second, cross-functional coopetition is also similar to market orientation in that both emphasize cooperation across functions or departments. In particular, a process view of market orientation suggests that knowledge about customers is disseminated across departments (e.g., Kohli and Jaworski 1990). However, coopetition and market orientation differ in an important way. Whereas market orientation focuses on promoting functional coordination and reducing interdepartmental conflict and friction (e.g., Narver and Slater 1990), cross-functional coopetition recognizes that interdepartmental conflict is not always unfavorable and can even produce specific benefits (Anderson and Narus 1990; Lado, Boyd, and Hanlon 1997).

The Effect of Cross-Functional Coopetition on Performance

We expect the joint occurrence of cross-functional cooperation and competition to enhance firm performance. The synergy of coopetition has been theorized to produce various benefits, including learning, cost savings, resource sharing, and innovation (Lado, Boyd, and Hanlon 1997). At an intraorganizational level, cooperation among competing units may be evident by absorbed and frequent interactions to uncover competing units' know-how (Tsai 2002). Therefore, we focus on the ability and intensity, respectively, of the nature of interdepartmental cooperation and examine the performance effects of cross-functional coopetition in more detail by investigating (1) the joint occurrence of cross-functional cooperative ability and competition and (2) the joint occurrence of cross-functional cooperative intensity and competition.

The joint occurrence of cross-functional cooperative ability and competition. Cross-functional cooperative ability refers to skills in recognizing the value of new market knowledge as well as assimilating, transforming, and deploying market knowledge across lateral cross-functional interactions, reflecting an absorptive capacity for cooperative knowledge transfer (Cohen and Levinthal 1990; Zahra and George 2002). The ability for cross-functional cooperation extends beyond the acquisition of market knowledge, because functional units that do not have the capacity for assimilation and transformation may not be able to internalize the available knowledge or to deploy it effectively.

Regarding the joint impact of cross-functional cooperative ability and competition on firm performance, it is possible that a negative effect may occur. In particular, Rindfleisch and Moorman (2003) show that a firm's customer orientation may deteriorate if the firm is engaged in a cooperative alliance with competitors; one reason for this decline is a low level of trust among alliance members. It is also noted that interfunctional rivalry decreases the level of trust in information (e.g., Maltz and Kohli 1996), which suggests that cross-functional competition coupled with cooperative ability could decrease a firm's customer-based performance.

However, we expect that the joint occurrence of cross-functional cooperative ability and competition has a positive effect on a firm's customer performance and financial performance. A combination of cross-functional cooperative ability and competition may nurture productive interactions (e.g., Brandenburger and Nalebuff 1996; Lado, Boyd, and Hanlon 1997; Uzzi 1999) that can facilitate internal efficiencies and sharing of best practice for successful deployment and exploitation of knowledge. In particular, cooperation among competitors can foster greater knowledge seeking and result in syncretic rents (Lado, Boyd, and Hanlon 1997). High cross-functional cooperative ability emphasizes the nature of gaining, absorbing, and sharing customer and market knowledge, whereas high cross-functional competition may provide departments with a strong incentive to share this knowledge to understand one another's positions better. Indeed, Tsai (2002) empirically shows that productive interactions can be nurtured when there is high competition for resources across a firm's SBUs because they are more likely to share information and exploit valuable knowledge stores. These productive interactions can also affect firm performance. For example, Von Hippel (1987) shows that cooperation among competing firms can improve each firm's profits.

As such, the joint effect of high cooperative ability and high competition may generate better problem solving in satisfying customer needs and higher performance (Hamel, Doz, and Prahalad 1989; Lado, Boyd, and Hanlon 1997; Tsai 2002). Therefore, we predict that at the functional level within the firm, the joint occurrence of cooperative ability and competition gives rise to better customer and financial performance.

> H_1: The joint occurrence of cross-functional cooperative ability and competition has a positive effect on a firm's (a) customer performance and (b) financial performance.

The joint effect of cross-functional cooperative intensity and competition. Cross-functional cooperative intensity refers to the degree to which lateral, interdepartmental interactions are frequent and close within an organization. Prior research shows that frequent cooperative interaction enhances the level of communication and knowledge sharing (Ghoshal, Korine, and Szulanski 1994; Tsai 2002). Frequent interactions across functional areas may also serve as a mechanism that helps functions both recognize where potentially useful market knowledge could serve and route it to the destinations across functional boundaries for improved outcomes (Huber 1991).

We expect that the joint occurrence of cross-functional cooperative intensity and competition has a positive influence on a firm's customer performance and financial performance. More specifically, frequent interactions can create more opportunities to share knowledge and ideas (Tsai 2002), and sharing information across competing functions can offer timely integration of knowledge within a firm (e.g., Uzzi 1997), which in turn stimulates superior performance. Intensive cooperation across competing departments may not only promote the development and conversion of tacit knowledge into a shared understanding of customer needs (Nonaka 1994; Rindfleisch and Moorman 2001) but also generate more collaborative and effective strategic decision making within the organization (Ghoshal, Korine, and Szulanski 1994). Indeed, a competitive interdepartmental structure coupled with frequent interactions may enable a firm to capitalize on the communication channels and ties embedded across its functional areas, thus leading to greater market and financial returns.

Furthermore, the joint occurrence of cross-functional cooperative intensity and competition may reduce knowledge overlap and create synergies. In particular, although the intraorganizational functions are likely to share similar resource constraints and market situations imposed by the firm, knowledge across departments is less likely to be redundant, given the traditional emphases across functions (Moorman and Rust 1999). With less redundancy, frequent interaction among competing departments provides access to novel and complementary market intelligence across functional knowledge silos (Granovetter 1973). This access to nonredundant information fosters better problem solving and decision making (e.g., Cummings 2004) and is essential for the creation of customer and financial value (e.g., Kohli and Jaworski 1990). Therefore, we predict that the joint occurrence of cross-functional cooperative intensity and competition enhances customer and financial performance.

> H_2: The joint occurrence of cross-functional cooperative intensity and competition has a positive effect on a firm's (a) customer performance and (b) financial performance.

The Mediating Role of Market Learning

We also argue that a firm with departments that simultaneously compete and cooperate may enhance its market learning, which in turn will generate higher performance. Market learning broadly refers to a firm's expertise in and knowledge stores of key marketing activities, such as developing new products, building brand image, responding to customers' needs, and establishing channel relationships (Menon et al. 1999; Moorman 1995; Srivastava, Shervani, and Fahey 1998). On the basis of the knowledge-based view of the firm, which posits that competitive advantage results from different knowledge stores and market expertise (e.g., Grant 1996; Nonaka 1994), we examine whether market learning is a route from which cross-functional coopetition enhances a firm's customer and financial performance.

The coexistence of cross-functional cooperation and competition is likely to create a synergy that improves a firm's market learning. First, the joint occurrence of cross-functional cooperative ability and competition offers the opportunity for customer and market knowledge to cross a firm's functional boundaries in an absorbed and effective way. With the facilitating effect of competition, brought about by the interest in knowing more about competitors' positions (Hamel, Doz, and Prahalad 1989; Tsai 2002), the cooperative ability to recognize, assimilate, and apply specialized market knowledge across competing departments is likely to enhance the ease and effectiveness with which learning may occur (Cohen and Levinthal 1989; Grant 1996; Szulanski 1996).

Second, the joint occurrence of cross-functional cooperative intensity and competition is also expected to improve a firm's market learning by enabling departments to enhance their knowledge stores in proactive and innovative ways. Intensive and frequent interactions can improve the transfer of complex knowledge (Hansen 1999) and adaptation to complex situations (Uzzi 1997). Furthermore, as we previously indicated, cross-functional competition can instill the motivation to understand other functions' positions. Thus, when intensive interactions occur across competing functions, this joint occurrence of cross-functional cooperative intensity and competition enables the competing departments to attain new and complementary knowledge that reduces the chance for interdepartmental misunderstanding of complex knowledge (Griffin and Hauser 1996) and offers the opportunity to diffuse novel ideas and promote greater learning (e.g., Ghoshal, Korine, and Szulanski 1994).

Therefore, we expect that the joint occurrence of cross-functional cooperative ability and competition and the joint occurrence of cross-functional cooperative intensity and competition enable a firm to generate stronger market learning.

> H_3: A firm's market learning is influenced by (a) the joint occurrence of its cross-functional cooperative ability and competition and (b) the joint occurrence of its cross-functional cooperative intensity and competition.

Extant literature in marketing shows a large and accumulating amount of evidence regarding the influence of organizational learning and market intelligence on firm performance (e.g., Kohli and Jaworski 1990; Maltz and Kohli 1996; Moorman, Zaltman, and Deshpandé 1992; Rindfleisch and Moorman 2001; Zhou, Yim, and Tse 2005). In addition, fundamental tenets of the knowledge-based view (Grant 1996; Kogut and Zander 1992) indicate that superior knowledge stores and organizational learning, as strategic assets, lead to improved financial and market performance. As previously argued, we believe that cross-functional coopetition (1) helps a firm build the necessary knowledge stores for superior market learning, in the form of the joint occurrence of cross-functional cooperative intensity and competition, and (2) enables the firm to exploit these knowledge stores for superior market learning, in the form of the joint occurrence of cross-functional cooperative ability and competition. In turn, superior market learning leads to improved firm financial and market performance (e.g., Dodgson 1993; Kogut and Zander 1992).

Therefore, we purport that the performance advantages of cross-functional coopetition are achieved through an underlying learning mechanism.

H_4: Market learning mediates the influence of the joint occurrence of cross-functional cooperative ability and competition on a firm's (a) customer performance and (b) financial performance.

H_5: Market learning mediates the influence of the joint occurrence of cross-functional cooperative intensity and competition on a firm's (a) customer performance and (b) financial performance.

Methodology

Sample

The sample consists of firms that operate in high-technology sectors such as biotechnology, software development, information technology, and electronics within China. As the second largest economy in the world, China represents a critical market to consider (Boisot and Child 1996; Walters and Samiee 2003). Using the official *China Basic Statistical Units Yearbook* as the sampling frame, we randomly selected 500 firms headquartered in metropolitan areas with nationwide operations in China.[2]

The data were collected from two key informants within the organization to minimize potential common method bias. In particular, common method bias concerns arise when both independent and dependent variables are measured by the same key informant (e.g., Van Bruggen, Lilien, and Kacker 2002). To specify the two key informants from each of the 500 firms, the firms were telephoned and requested to select one midlevel manager randomly (e.g., sales, marketing, research-and-development [R&D] department manager) and one top executive (e.g., chief executive officer, general manager). The key informants were carefully chosen to ensure that they had the knowledge and background to complete the questionnaire in a thoughtful manner. The midlevel managers and senior executives were interviewed separately (Atuahene-Gima and Li 2002). In generating these responses for each company, top executives answered survey questions about organizational performance, market learning, and firm demographic information, whereas midlevel department managers answered questions about cross-functional cooperative forces, competitive forces, and some items of firm performance as a check of data validity.

A total of 329 firms (658 managers) met the screening criteria (i.e., the knowledge and background to respond and two key informants from each firm) and initially agreed to participate. After matching key informants and deleting missing data, we ended up with usable information from 163 firms (326 informants). As a result, we obtained a response rate of 32.6% of the original sample (163/500) and 49.5% of the qualified informants (163/329). To test for nonresponse bias, we compared the demographic data of 70 firms that agreed to participate but did not complete the interviews with those of the completed and useful responses. We found no significant difference in firm size, industry, location, and ownership.

The results testing the performance items that were designated to check informant response validity (sales growth, customer satisfaction, and return on investment [ROI]) exhibit high consistency between top executives and midlevel managers (Guttmans: .83, .88, and .87, respectively). To ensure that the senior executives were appropriate and reliable, we examined the extent of their involvement in making strategic decisions in the firm (Li and Atuahene-Gima 2001); the results show high strategic involvement (M = 6.81 on a seven-point scale). Among the midlevel managers, several functional areas were covered, including sales (38%), marketing (33%), R&D (21%), and others (8%). Note that all functional areas may generate market knowledge to enhance customer value (Kohli and Jaworski 1990; Narver and Slater 1990) and that the process of market intelligence dissemination across functional boundaries can be evaluated by midlevel managers from marketing and other functional areas (e.g., Fisher, Maltz, and Jaworski 1997; Maltz and Kohli 1996). In general, 68% of informants had more than five years' experience in the firm, and 91% had at least two years' experience with the firm. Thus, the key informants were reasonably knowledgeable about the cross-functional dynamics within the firm (e.g., Johnson, Sohi, and Grewal 2004). Among the firms in our sample, 72.5% were small or medium in size, with 500 or fewer employees. The largest industry segment was electronics (36.3%), followed by information technology (28.1%) and biotechnology (18.7%).

Measures

We assessed all constructs as multi-item measures. In addition, we scored each item on a seven-point scale, ranging from "strongly disagree" (1) to "strongly agree" (7). The questionnaire was prepared in English, translated into Chinese by independent translators, and then back translated into English to ensure accuracy and follow appropriate guidelines (Brislin 1970). The questionnaire was then pretested with 17 Chinese managers in a pilot study. Reponses from the pretests ensured the relevance and equivalence of the final measures. We describe all measures in Table 1 and detail them in the Appendix. We discuss the measurement results in the "Results" section.

We measured *cross-functional cooperative intensity* with six items that we constructed to assess the degree to which the lateral cross-functional interactions are frequent and close. We adapted this measure in part from similar research in marketing that focuses on relational embedded-

---

[2] These 500 firms were equally selected from five major cities, including Shanghai in the East, Chengdu in the West, Beijing in the North, Guangzhou in the South, and Changsha in Central China. We selected this equal distribution to minimize bias of regional market heterogeneity.

TABLE 1
Main Constructs, Sources, and Definitions

Construct	Construct Sources	Construct Definition	Key Informants in the Dyad
Cross-functional cooperative intensity	Antia and Frazier (2001); Rindfleisch and Moorman (2001)	The extent of the frequency and closeness of the lateral social interactions among functional areas within the firm.	Midlevel managers
Cross-functional cooperative ability	Cohen and Levinthal (1990); Szulanski (1996); Zahra and George (2002)	The ability to assimilate and deploy market knowledge in lateral interactions among functional areas.	Midlevel managers
Cross-functional competition	Levitt (1969); Houston et al. (2001); Maltz and Kohli (1996); Ruekert and Walker (1987)	The degree to which departments compete both for limited tangible and intangible resources and for strategic importance, power, and department charter.	Midlevel managers
Market learning	Menon et al. (1999); Moorman (1995); Srivastava, Shervani, and Fahey (1998)	The firm's expertise in and knowledge stores of key marketing activities (i.e., developing new product development, building brand image, sensemaking customers' current and potential needs, and others).	Top executives
Firm performance	Moorman and Rust (1999)	The perceptions of strategic market performance and financial performance (i.e., market share growth, customer loyalty, new product development, sales growth, customer satisfaction, selling costs, customer retention, customer lifetime value, and ROI).	Top executives

ness at other organizational levels (e.g., Antia and Frazier 2001; Rindfleisch and Moorman 2001). Our *cross-functional cooperative ability* measure includes six items that examine the core underlying ability to evaluate, assimilate, and exploit market knowledge transferred from other departments during cross-functional interactions. In line with the process perspective of absorptive capacity (Cohen and Levinthal 1990; Lane, Salk, and Lyles 2001), this measure extended the three-item scale that Szulanski (1996) initially suggested and validated.

The measure of *cross-functional competition* included 11 items that assessed the extent to which functional areas compete with one another for tangible and intangible resources, strategic importance, and charter (Ruekert and Walker 1987) and the extent to which departments struggle with one another because of divergent departmental goals and strategic priorities (Houston et al. 2001; Levitt 1969; Maltz and Kohli 1996). We dropped one item to enhance scale validity (see the Appendix), which resulted in a ten-item measure.

We developed the ten-item measure of *market learning* to capture the firm's expertise in and knowledge stores of key marketing activities, adapting our measure from Menon and colleagues' (1999) list of key marketing activities. With these ten items, we model market learning as a higher-order construct of three subdimensions. This approach is similar to extant literature that uses a second-order factor to examine cooperative competency (Sivadas and Dwyer 2002). Our three subdimensions include market learning in product/service (i.e., developing a new product/service, promoting and selling a product/service, and refining and repositioning an existing product/service), customer (i.e., building a brand image among customers, developing extensive customer service capabilities, and making sense of customers' current and potential needs), and channel (i.e., pricing below competitors, establishing marketing and distribution networks in the market, collaborating with business partners, and knowing about business environment). The results show good overall fit: $\chi^2 = 36.85$, d.f. = 32, $p > .05$; comparative fit index (CFI) = .981, goodness-of-fit index (GFI) = .957, and root mean square error of approximation (RMSEA) = .031; as we expected, the three first-order latent constructs loaded significantly on market learning (product/services: $\lambda = .753, p < .01$; customer: $\lambda = .816, p < .01$; channel: $\lambda = .610, p < .01$).

In terms of performance, we examined two broad metrics of *firm performance*. To assess the firm's strategic market and financial outcomes, such as market growth and ROI, we used a four-item measure of financial performance that we adapted, in part, from the work of Moorman and Rust (1999). In addition, to assess the firm's customer-based performance, including its customer satisfaction and customer retention, we used a four-item measure of customer performance that we also adapted, in part, from the work of Moorman and Rust (1999).

To control for any extraneous effects due to firm or industry factors, we also included several covariates. In particular, we included firm size and firm ownership to control for economies of scale and resource capacity. We measured *firm size* using the number of employees, which is consistent with extant research (e.g., Lee and Grewal 2004; Luo 2001), and we measured *firm ownership* using a dummy variable (1 = firms with foreign ownership, 0 = otherwise [i.e., state- or private-owned firms]) due to strategic and operations differences between foreign and local firms in China (e.g., Peng and Luo 2000). In addition, we included two industry variables to control for market effects that have been reported to influence a firm's strategic choices and performance outcomes (e.g., Atuahene-Gima and Li 2002; Li and Calantone 1998; Zhou, Yim, and Tse 2005). Specifically, we included *industry hostility*, which we measured using the five-item scale that Jaworski and Kohli (1993) developed, and *market volatility*, which we measured using a four-item scale that is well-accepted in marketing literature (e.g., Li and Calantone 1998; Slater and Narver 1994; Zhou, Yim, and Tse 2005).

Analysis and Results

Measure Validation

Following the work of Anderson and Gerbing (1988), we test the validity of the measures using confirmatory factor analysis. Overall model statistics indicate that the chi-square for the measurement model with all constructs is 1427.18 ($p = .00$), with an acceptable ratio of chi-square to degrees of freedom of 1.67; the CFI, GFI, adjusted goodness-of-fit index (AGFI), and RMSEA are .96, .94, .93, and .05, respectively.

Convergent validity for each of the measures, as indicated by Cronbach's alpha, all exceeded the .7 benchmark (Nunnally 1978). As Table 2 shows, the minimum reliability across the measures is .90. Furthermore, at .71 or higher, the average variance extracted (AVE) for each construct exceeds the .5 benchmark (Fornell and Larcker 1981).

Discriminant validity of the measures was supported with two different approaches. First, we conducted a series of chi-square difference tests by examining pairs of constructs using two models (one constrained model and one unconstrained model) for each pair of measured constructs (Anderson and Gerbing 1988). The results suggest that unconstrained models fit the data better than constrained models, indicating discriminant validity. Second, we compared the AVE of the individual constructs with the squared correlation between construct pairs (Fornell and Larcker 1981). As is evident in Table 2, the largest correlation is that between customer performance and financial performance ($r = .58$), which, when squared ($r^2 = .336$), is less than the AVE for either customer performance or financial performance (AVE = .75 and .72, respectively). In all cases, we find that the AVEs exceed the squared correlations, which again confirms discriminant validity.

The Effects of Cross-Functional Coopetition

Using regression analysis, we capture the effects of cross-functional coopetition with the interaction between the cooperation and competition constructs. We mean-centered these variables before creating the interaction to reduce any collinearity between the main and the interaction effects (Aiken and West 1996). Following Aiken and West's (1996) and Cohen and Cohen's (1983) recommendations, we entered the variables into the model in distinct steps: We entered the covariates, then the main effects, and finally the interaction effects. To examine the incremental effect of including the interaction between cross-functional cooperation and competition, we examined the change in R-square (ΔR^2), which indicates the value of including additional variables in a model (Cohen and Cohen 1983) and the value of interactions in particular (Aiken and West 1996). The results for the effects of cross-functional coopetition on firm performance appear in Table 3.

The results show that the interaction between (i.e., the joint occurrence of) cross-functional cooperative ability and competition has a positive and significant effect on customer performance ($b = .19$, $p < .01$) and on financial performance ($b = .26$, $p < .01$), in support of H_1 (see Table 3). In addition, the results show that the joint occurrence of cross-functional cooperative intensity and competition has a positive and significant effect on customer performance ($b = .23$, $p < .01$) and on financial performance ($b = .16$, $p < .05$), in support of H_2 (see also Table 3). Moreover, entry of the joint occurrence of cross-functional competition and cooperative ability and of the joint occurrence of cross-functional competition and cooperative intensity in the third step explained a significant level of additional variance beyond the main effects. Specifically, these two types of cross-functional coopetition explained 8% more variance for customer performance ($F_{(2, 159)} = 3.96$, $p < .01$) and 10% more variance for financial performance ($F_{(2, 159)} = 4.43$, $p < .01$).

For our examination of whether there are any differential effects for the two types of organizational performance, model comparisons show that the joint occurrence of cross-

TABLE 2
Descriptive Statistics of Key Constructs

Variables	M	SD	AVE	CI	CA	CM	LEARN	CUPF	FIPF
Cross-functional cooperative intensity (CI)	5.28	1.23	.74	.91					
Cross-functional cooperative ability (CA)	5.89	1.14	.73	.43**	.91				
Cross-functional competition (CM)	4.19	1.31	.71	.32**	.36**	.90			
Market learning (LEARN)	5.10	1.18	.72	.18*	.37**	–.12	.90		
Customer performance (CUPF)	5.41	1.05	.75	.36**	.24**	–.17*	.39**	.95	
Financial performance (FIPF)	5.08	1.36	.72	.27**	.35**	–.10	.37**	.58**	.90

*$p <$ 05.
**$p < .01$.
Notes The coefficient alpha for each measure is on the diagonal, and the intercorrelations are on the off-diagonal

TABLE 3
The Effect of Cross-Functional Coopetition on Customer and Financial Performance

Independent Variables	Customer Performance			Financial Performance		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Coopetition Effects						
Joint occurrence of cross-functional cooperative ability and competition (CA × CM)			.19**			.26**
Joint occurrence of cross-functional cooperative intensity and competition (CI × CM)			.23**			.16*
Main Effects						
Cross-functional cooperative ability (CA)		.27**	.29**		.36**	.39**
Cross-functional cooperative intensity (CI)		.32**	.36**		.22**	.23**
Cross-functional competition (CM)		–.16*	–.13		–.11	–.10
Covariates						
Firm size	.06	.04	.04	.16*	.17*	.16*
Firm ownership	.16*	.16*	.14	.09	.08	.08
Industry hostility	.24**	.23**	.24**	.22**	.21**	.21**
Market volatility	–.09	–.11	–.10	–.12	–.16*	–.16
Overall Model						
F value	8.83**	10.65**	8.87**	7.21**	11.77**	10.04**
Adjusted R^2	.25	.43	.51	.21	.45	.55
ΔR^2		.18	.08		.24	.10
Incremental F		9.52**	3.96**		13.47**	4.43**

*$p < .05$.
**$p < .01$

functional cooperative ability and competition has a stronger effect on financial performance than on customer performance ($\chi^2_{diff} = 11.28$, d.f.$_{diff} = 1$, $p < .01$) and that the joint occurrence of cross-functional cooperative intensity and competition has a stronger effect on customer performance than on financial performance ($\chi^2_{diff} = 3.71$, d.f.$_{diff} = 1$, $p < .10$). It is possible that frequent interaction among competing departments helps build customer retention and loyalty, whereas improvements in a firm's financial performance require more absorptive and assimilating cooperation among competing departments.

The Mediating Role of Market Learning

To examine the extent to which market learning mediates the effects of cross-functional coopetition, we relied on the three-step mediated regression approach that Baron and Kenny (1986) recommend. More specifically, to establish mediation of market learning, coopetition must affect market learning and performance, and marketing learning must affect performance. As we previously indicated, the results show that cross-functional coopetition affects performance (see Models 3 and 6 in Table 3). Therefore, to establish mediation, cross-functional coopetition must affect market learning, and market learning must affect performance. The results show that the joint occurrence of cross-functional cooperative ability and competition has a positive effect on market learning (b = .22, $p < .01$) and that the joint occurrence of cross-functional cooperative intensity and competition has a positive effect on market learning (b = .17, $p < .05$), as Model 7 shows (see Table 4). These results support H_3. For our examination of the effect of market learning on performance, the results indicate that market learning has a positive and significant effect on a firm's customer performance (b = .32, $p < .01$) and financial performance (b = .28, $p < .01$).[3]

Furthermore, as Table 4 shows, inclusion of market learning in the model reduces the strength of the effect of cross-functional coopetition on performance. In particular, the joint effect of cross-functional cooperative ability and competition on customer performance and the joint effects of cross-functional cooperative intensity and competition on both customer performance and financial performance are no longer significant (see Models 8 and 9 in Table 4). Thus, market learning fully mediates the joint effect of cross-functional cooperative ability and competition on customer performance, fully mediates the joint effect of cross-functional cooperative intensity and competition on customer performance, and fully mediates the joint effect of cross-functional cooperative intensity and competition on

---

[3] Although common variance between market learning and firm performance may influence this effect, we examined whether the relationship between market learning and firm performance would differ when firm performance was reported by midlevel managers rather than by senior executives. The results indicate that there remains a strong relationship between market learning (reported by senior executives) and firm performance (reported by midlevel managers): r = 31, $p < .01$.

TABLE 4
The Mediating Effect of Market Learning

Independent Variables	Model 7: Market Learning	Model 8: Customer Performance	Model 9: Financial Performance
Mediating Effect			
Market learning		.32**	.28**
Coopetition Effects			
Joint occurrence of cross-functional cooperative ability and competition (CA × CM)	.22**	.11	.20**
Joint occurrence of cross-functional cooperative intensity and competition (CI × CM)	.17*	.13	.12
Main Effects			
Cross-functional cooperative ability (CA)	.37**	.20**	.26**
Cross-functional cooperative intensity (CI)	.29**	.24**	.18**
Cross-functional competition (CM)	–.12	–.09	–.09
Covariates			
Firm size	.05	.09*	.16*
Firm ownership	.10	.10	.06
Industry hostility	.24**	.18**	.20*
Market volatility	–.10	–.11*	–.17
Overall Model			
F value	14.87**	10.92**	9.62**
R^2	.53	.60	.58

*$p <$ 05.
**$p < .01$.

financial performance. Moreover, although the interaction effect of cross-functional cooperative ability and competition on financial performance remains significant, its effect diminishes (see Table 4), illustrating partial mediation. These results offer strong support for H_4 and H_5 and the argument that cross-functional coopetition realizes performance advantages through market learning.

Model Comparisons

To examine the extent to which our results are robust across different models, we used structural equation modeling (SEM) to compare alternative models. In particular, we fit the data with three models (Morgan and Hunt 1994; Zhou, Yim, and Tse 2005). In the first model (SEM1), all variables directly influence the two types of firm performance, and we do not include any mediation effect of market learning. The second model (SEM2) involves full mediation of market learning (i.e., the joint occurrence of cross-functional cooperative ability and competition and the joint occurrence of cross-functional cooperative intensity and competition are not modeled to affect performance). The third model (SEM3) allows partial mediation of market learning. All three models include the aforementioned covariates.

The results indicate that fit indexes for SEM1 are $\chi^2 = 1749.93$, d.f. = 852, $p = .00$; CFI = .90, GFI = .88, AGFI = .87, and RMSEA = .06; fit indexes for SEM2 are $\chi^2 = 1602.98$, d.f. = 855, $p = .00$; CFI = .92, GFI = .91, AGFI = .89, and RMSEA = .06; and fit indexes for SEM3 are $\chi^2 = 1578.54$, d.f. = 851, $p = .00$; CFI = .94, GFI = .92, AGFI = 90, and RMSEA = .05. Comparison across these three alternative SEM models illustrates that SEM3, with partial mediation effects, fits the data best, fitting the data better than SEM1 with no meditation effects ($\chi^2_{\text{diff}} = 171.39$, d.f.$_{\text{diff}} = 1$, $p < .01$) and better than SEM2 with full mediation effects ($\chi^2_{\text{diff}} = 24.44$, d.f.$_{\text{diff}} = 4$, $p < .01$). In other words, these findings show that market learning partially mediates the influence of cross-functional coopetition effects on firm performance, consistent with the regression results in Table 4. Furthermore, the structural paths in SEM3 appear to be consistent with the beta coefficients from the regression results. Specifically, SEM3 results show support for the market learning–customer performance (b = .35, $p < .01$) and market learning–financial performance (b = .29, $p < .01$) linkages,[4] comparable to the counterpart beta coefficients of .32 ($p < .05$) and .28 ($p < .05$) regression results in Table 4.[5] Overall, the results for SEM confirm our conceptual framework, demonstrating the performance benefits of cross-functional coopetition and the role of market learning as a mediator of these benefits.

---

[4] We assessed SEM path coefficients significance through a bootstrapping procedure with 1000 resamples.

[5] Although we found a significant path from cross-functional coopetive intensity to customer performance in SEM3 (b = .18, $p < .05$), which was not significant in the regression analysis, specific differences between regression results and SEM results may be due to different estimation approaches (ordinary least squares in the regression approach and maximum likelihood in the SEM approach) and to the accounting for measurement error in SEM.

Discussion

The recently emerging research stream on coopetition has been largely limited to the interfirm level (e.g., Rindfleisch and Moorman 2003; Zeng and Chen 2003). Only a few studies have examined coopetition at the intraorganizational level, focusing on SBU-level knowledge sharing in general (Tsai 2002) or within a multinational organization (e.g., Ghoshal, Korine, and Szulanski 1994). Coopetition at the cross-functional level has largely been ignored in both the marketing and the management literature. Notably, Rindfleisch and Moorman (2001) point out that the marketing literature has devoted surprisingly little systematic research on interfirm coopetition, let alone intrafirm coopetition.

Our focus on the simultaneous role of cooperation and competition in cross-functional relationships advances the recent emphasis on market knowledge transfer across functional boundaries for superior learning and performance (Maltz and Kohli 1996, 2000). We argue that because cross-functional cooperation involves the ability for and the intensity of knowledge transfer and because competition can facilitate knowledge transfer given an underlying incentive to understand competing functions' positions, the joint occurrence of cross-functional cooperation and competition can create synergies that translate into competitive advantage. Overall, our results indicate that cross-functional coopetition has an important impact on a firm's customer-based and financial-based performance and that market learning mediates these performance returns. Thus, although a firm's functional areas may experience competition for resources and strategic emphasis, cooperative forces are necessary to shape not only its market learning but also its customer and financial performance. Consequently, our findings regarding the synergy of simultaneous cooperation and competition at the functional level provide some empirical evidence of the "positive-sum" or syncretic rents of interfirm coopetition that Lado, Boyd, and Hanlon (1997) theorize.

Importantly, our results also show that cross-functional coopetition can affect firm performance in different ways. For example, the joint occurrence of cross-functional cooperative intensity and competition has a stronger positive effect on customer-based performance than on financial-based performance, potentially because frequent interactions among competing departments may not be an efficient approach (Lado, Boyd, and Hanlon 1997); yet the communication offers additional information that can benefit consumers (Fisher, Maltz, and Jaworski 1997; Griffin and Hauser 1996). Our introduction of two important concepts of cross-functional cooperation—namely, cooperative intensity and cooperative ability—also offers important insight for intrafirm knowledge transfer and performance outcomes. A combination of the two cooperation concepts in cross-functional coopetition may help solve the "fail-to-use" problem of sticky yet valuable market knowledge (Maltz and Kohli 1996), such that the joint occurrence of cross-functional cooperative ability and competition offers the ability to recognize and assimilate valuable yet tacit market knowledge and the joint occurrence of cross-functional cooperative intensity and competition provides the channels for knowledge transfer across functional boundaries within the firm.

Although prior work has suggested a link between coopetition and performance (e.g., Brandenburger and Nalebuff 1996; Hamel, Doz, and Prahalad 1989; Lado, Boyd, and Hanlon 1997), few studies have explored the reasons for this link. Drawing from the knowledge-based view of the firm, we probe for an underlying learning mechanism and point out that coopetition at the intraorganizational level influences firm performance through market learning. Specifically, our results indicate that market learning plays a major mediating role in the extent to which cross-functional coopetition influences firm performance. Thus, the value of simultaneous cooperation and competition within cross-functional interactions lies in how they influence a firm's market learning, which in turn affects the firm's performance. This finding highlights a critical implication for marketing scholars because it uncovers an important role of marketing knowledge and intraorganizational learning in exploiting cooperative ability and intensity among competing departments for better firm performance.

Managerial Implications

Our research also offers insight to managers regarding the importance of simultaneously managing cooperation and competition in cross-functional interactions. Although executives may recognize that market knowledge transfer across functional areas is valuable, they must also understand that it is a complicated process due, in part, to the underlying nature of cooperation and competition in cross-functional relationships.

Notably, our results regarding the strong, beneficial influence of cross-functional coopetition on firm performance indicate that fostering cooperation but quenching competition may limit a firm's full performance potential. In other words, our research directs managers to understand that competition and conflicts among marketing, R&D, finance, and other departments are not always harmful. Instead, fostering the joint occurrence of cross-functional cooperation and competition can direct conflicts toward constructive interactions and actually promote a firm's overall market learning and performance. For example, Royal Dutch Shell had two functions that were focused on bettering their own individual departments and competing for business within the organization. When these competing functions came together to share knowledge, the organization reaped higher performance (Burress and Wallace 2003). This example, coupled with our research findings, informs managers that a firm should attempt to cooperatively bridge the competitive nature of its functional areas and leverage the joint occurrence of competition and cooperation for improved firm customer and financial performance.

Indeed, the overwhelming push for lean budgets and downsizing in today's business practice is likely to generate high levels of competition among functional areas. This competition can often be the result of how resources are apportioned across departments. For example, firms frequently cut marketing budgets to save costs, often because marketing expenditures are not easily tied to financial outcomes (Rust et al. 2004). As a result, with inherent and often unavoidable competitive pressures between marketing and other functions (i.e., fighting for more budget from the shrinking sum within a firm), executives should consider specific approaches that can instill cooperative ability and intensity across the competing departments, such as use of cross-functional teams to focus on organizational goals rather than purely functional goals (Griffin and Hauser 1996; Maltz and Kohli 2000), implementation of explicit employee incentives that focus on the firm's long-term objectives (Hauser, Simester, and Wernerfelt 1994), illustration of marketing expenditure effects on shareholder value to increase the credibility of the marketing function within the firm (Rust et al. 2004), and structural changes to offer more opportunities for informal interactions among departments within the firm.

Our findings also point managers to the valuable mediating role of market learning in the process of converting cross-functional coopetition into higher performance. Specifically, our results indicate that engaging in more intensive and absorbed collaborations across competing functional boundaries can enhance a firm's overall performance through a learning route. Thus, managers should nurture the frequency of interactions among competing departments and the absorptive capacity for cooperation among competing functions to promote intraorganizational learning, which ultimately advances firm performance.

Finally, managers who want to enhance organizational learning and are interested in more readily transferring knowledge within the firm should understand that the culprit of this problem may result not simply from low intensity of cross-functional cooperation but also from a lack of assimilation and deployment of market knowledge. Overall, both cooperative intensity and ability may help functional areas recognize where market knowledge resides within the organization, assimilate such knowledge, and then deploy this market knowledge for improved learning and subsequent superior firm performance.

Further Research and Current Limitations

First, further research on cross-functional coopetition could examine differences across functional areas regarding the cooperative ability to recognize and assimilate market knowledge. In particular, our research unveiled that departments are not the same in their abilities to assimilate and use market knowledge that is available within the firm. Thus, further research should pragmatically recognize the difference among functional areas regarding their abilities to absorb market knowledge and explore antecedents of this difference.

Second, further research could track how learning mediates the relationship between cross-functional coopetition and organizational performance over time. Our research indicated that market learning plays a mediating role in transferring cross-functional coopetition into customer and financial performance. However, it would be interesting to examine how cross-functional level learning may translate over time to build stronger cross-functional cooperative intensity and ability. Overall, the notion of cross-functional coopetition may provide future marketing research a novel way of thinking about the complexities of market knowledge transfer across functional boundaries within the firm.

Appendix
Operationalization of Measures

Cross-Functional Cooperative Intensity (6-Item Measure)

New scale based on the work of Antia and Frazier (2001), Narver and Slater (1990), and Rindfleisch and Moorman (2001): reliability = .91 and AVE = .74 (reported by midlevel managers).

1. Departments here share communications frequently in our business.
2. All departments frequently discuss common problems in our business.
3. Marketing personnel share close ties with people in other departments.
4. Our relationship with other departments is mutually gratifying and highly cohesive.
5. We expect that our strong interdepartmental social relationship will exist far into the future.
6. There is little informal interaction among people from different departments. (Reverse coding)

Cross-Functional Cooperative Ability (6-Item Measure)

New scale based in part on the work of Szulanski (1996) and Cohen and Levinthal (1990): reliability = .91 and AVE = .73 (reported by midlevel managers).

During the interdepartmental interactions, our departments here have strong abilities to

1. Identify new and useful market knowledge transferred from other departments.
2. Understand new and useful market knowledge transferred from other departments.
3. Value new and useful market knowledge transferred from other departments.
4. Assimilate new and useful market knowledge transferred from other departments.
5. Apply new and useful market knowledge transferred from other departments.
6. Exploit new and useful market knowledge transferred from other departments.

Cross-Functional Competition (11-Item Measure)

New scale based on the work of Houston and colleagues (2001), Maltz and Kohli (1996), and Ruekert and Walker (1987): reliability = .90 and AVE = .71 (reported by midlevel managers).

1. We regularly compete for the limited resources across departments
2. When members of several departments talk about distribution of resources (i.e., capital, personnel) across departments, tensions frequently run high

3. Functional areas regularly compete with each other for more mental attention and time from top executives.
4. To get more resources for our department, other departments oftentimes have to make sacrifices.
5. Individual departments here try to obtain more time and attention from senior managers even at the costs of other functions.
6. Each department is constantly compared and benchmarked with other departments to improve efficiency in the organization.
7. Most departments here try to gain more strategic importance and power inside the firm.
8. The objectives pursued by the marketing department are incompatible with those of other departments (e.g., manufacturing, IT [information technology], operations). (Item dropped to enhance scale validity)
9. Protecting one's departmental turf is considered to be a way of life in this business.
10. Individual departments here tend to outperform others for a better department charter.
11. People from different departments feel that the goals of their respective departments are in harmony with each other. (Reverse coding)

Market Learning (10-Item Measure)

New scale based on the work of Menon and colleagues (1999) and Srivastava, Shervani, and Fahey (1998): reliability = .90 and AVE = .72 (reported by top executives).

Regarding market knowledge, compared to other companies in your industry, my company has a strong position in terms of the expertise in

1. Developing new product/service.
2. Promoting and selling product/service.
3. Refining and repositioning existing product/service.
4. Building brand image among customers.
5. Developing extensive customer service capabilities.
6. Sensemaking customers' current and potential needs.
7. Pricing below competitors.
8. Establishing marketing and distribution networks in the market.
9. Collaborating with business partners.
10. Knowing about business environment (tax, labor, institutional regulations, etc.).

Organizational Performance

Customer Performance (4-Item Measure)

Adapted in part from the work of Moorman and Rust (1999) and Narver and Slater (1990): reliability = .95 and AVE = .75 (reported by top executives).

1. Customer loyalty.
2. Customer satisfaction.
3. Customer lifetime value.
4. Customer retention.

Financial Performance (4-Item Measure)

Adapted in part from the work of Moorman and Rust (1999) and Narver and Slater (1990): reliability = .90 and AVE = .72 (reported by top executives).

1. Market share growth.
2. Sales growth.
3. Reducing selling costs.
4. ROI.

REFERENCES

Aiken, Leona and Stephen G. West (1996), *Multiple Regression. Testing and Interpreting Interactions*. Thousand Oaks, CA: Sage Publications.

Anderson, Erin and Barton Weitz (1989), "Determinants of Continuity in Conventional Industrial Channel Dyads," *Marketing Science*, 8 (4), 310–23.

Anderson, James C. and David W. Gerbing (1988), "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin*, 103 (3), 411–23.

——— and James A. Narus (1990), "A Model of Distributor Firm and Manufacturer Firm Working Partnerships," *Journal of Marketing*, 54 (January), 42–58.

Antia, Kersi and Gary L. Frazier (2001), "The Severity of Contract Enforcement in Interfirm Channel Relationships," *Journal of Marketing*, 65 (October), 67–81.

Atuahene-Gima, Kwaku and Haiyang Li (2002), "When Does Trust Matter? Antecedents and Contingent Effects of Supervisee Trust on Performance in Selling New Products in China and the United States," *Journal of Marketing*, 66 (July), 61–81.

Baron, Reuben M. and David A. Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (6), 1173–82.

Boisot, Max and John Child (1996), "From Fiefs to Clans and Network Capitalism: Explaining China's Emerging Economic Order," *Administrative Science Quarterly*, 41 (3), 600–628.

Brandenburger, Adam and Barry Nalebuff (1996), *Co-opetition*. New York: Currency Doubleday.

Brislin, R.W. (1970), "Back Translation for Cross-Cultural Research," *Journal of Cross-Cultural Psychology*, 1 (3), 185–216.

Burress, Ann and Sharon Wallace (2003), "Brainstorming Across Boundaries at Shell," *KM Review*, 6 (2), 20–23.

Burt, Ronald S. (1992), *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.

Cohen, Jacob and Patricia Cohen (1983), *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 2d ed. Hillsdale, NJ. Lawrence Erlbaum Associates.

Cohen, Wesley M. and Daniel A. Levinthal (1989), "Innovation and Learning: The Two Faces of R&D," *The Economic Journal*, 99 (September), 569–96.

——— and ——— (1990), "Absorptive Capacity: A New Perspective on Learning and Innovation," *Administrative Science Quarterly*, 35 (March), 128–52.

Cummings, Jonathon N. (2004), "Work Groups, Structural Diversity, and Knowledge Sharing in a Global Organization," *Management Science*, 50 (3), 352–64.

Deshpandé, Rohit and Gerald Zaltman (1982), "Factors Affecting the Use of Market Research Information: A Path Analysis," *Journal of Marketing Research*, 19 (February), 14–31.

Dodgson, Mark (1993), "Organizational Learning: A Review of Some Literatures," *Organization Studies*, 14 (3), 375–94.

Fisher, Robert J., Elliot Maltz, and Bernard J. Jaworski (1997), "Enhancing Communication Between Marketing and Engineering: The Moderating Role of Relative Functional Identification," *Journal of Marketing*, 61 (July), 54–70.

Fornell, Claes and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 19 (February), 39–50.

Frankwick, Gary, James Ward, Michael Hult, and Peter Reingen (1994), "Evolving Patterns of Organizational Beliefs in the Formation of Strategy," *Journal of Marketing*, 58 (April), 96–110.

Ghoshal, Sumantra, Harry Korine, and Gabriel Szulanski (1994), "Interunit Communication in Multinational Corporations," *Management Science*, 40 (1), 96–110.

Gilmour, David (2003), "How to Fix Knowledge Management," *Harvard Business Review*, 81 (October), 16–17

Granovetter, Mark (1973), "The Strength of Weak Ties," *American Journal of Sociology*, 78 (6), 1360–80.

——— (1985), "Economic Actions and Social Structure: The Problem of Embeddedness," *American Journal of Sociology*, 91 (3), 481–510.

Grant, Robert M. (1996), "Prospering in Dynamically-Competitive Environments: Organizational Capability as Knowledge Integration," *Organization Science*, 7,(4), 375-387.

Gray, Peter H. and Darren B. Meister (2004), "Knowledge Sourcing Effectiveness," *Management Science*, 50 (6), 821–34.

Griffin, Abbie and John R. Hauser (1992), "Patterns of Communication Among Marketing, Engineering, and Manufacturing: A Comparison Between Two New Product Teams," *Management Science*, 38 (3), 360–73.

——— and ——— (1996), "Integrating R&D and Marketing: A Review and Analysis of the Literature," *Journal of Product Innovation Management*, 13 (3), 191–215.

Gulati, Ranjay (1998), "Alliances and Networks," *Strategic Management Journal*, 19 (4), 293–317.

Gupta, Anil K. and Vijay Govindarajan (2000), "Knowledge Flows Within Multinational Corporations," *Strategic Management Journal*, 21 (4), 473–95.

Hamel, Gary, Yves L. Doz, and C.K. Prahalad (1989), "Collaborate with Your Competitors and Win," *Harvard Business Review*, 67 (1), 133–39.

Hammond, Janice H. (1994), "Barilla SpA (A)," Harvard Business School Case No. 9-694-046.

Hansen, Morten T. (1999), "The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge Across Organization Subunits," *Administrative Science Quarterly*, 44 (1), 82–111.

Hauser, John R., Duncan I. Simester, and Birger Wernerfelt (1994), "Customer Satisfaction Incentives," *Marketing Science*, 13 (4), 327–50

Heide, Jan B. and Anne S. Miner (1992), "The Shadow of the Future: Effects of Anticipated Interaction and Frequency of Contact on Buyer-Seller Cooperation," *Academy of Management Journal*, 35 (2), 265–91

Homburg, Christian, John P. Workman Jr., and Harley Krohmer (1999), "Marketing's Influence Within the Firm," *Journal of Marketing*, 63 (April), 1–17.

Houston, Mark B., Beth A. Walker, Michael D. Hutt, and Peter H. Reingen (2001), "Cross-Unit Competition for a Market Charter: The Enduring Influence of Structure," *Journal of Marketing*, 65 (April), 19–34.

Huber, George P. (1991), "Organizational Learning: The Contributing Processes and the Literature," *Organization Science*, 2 (February), 88–115.

Jaworski, Bernard J. and Ajay K Kohli (1993), "Market Orientation: Antecedents and Consequences," *Journal of Marketing*, 57 (July), 53–70.

Johnson, Jean, Ravipreet Sohi, and Rajdeep Grewal (2004), "The Role of Relational Knowledge Stores in Interfirm Partnering," *Journal of Marketing*, 68 (July), 21–36.

Kerwin, K. and D. Woodruff (1992), "Can Olds Hitch Its Wagon to Saturn's Star?" *BusinessWeek*, (November 23), 74.

Kogut, Bruce and Udo Zander (1992), "Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology," *Organization Science*, 3 (3), 383–97

Kohli, Ajay K. and Bernard J. Jaworski (1990), "Market Orientation: The Construct, Research Propositions, and Managerial Implications," *Journal of Marketing*, 54 (April), 1–18.

Lado, Augustin A., Nancy G. Boyd, and Susan C. Hanlon (1997), "Competition, Cooperation, and the Search for Economic Rents: A Syncretic Model," *Academy of Management Review*, 22 (1), 110–41.

Lane, Peter J., Jane E Salk, and Marjorie A. Lyles (2001), "Absorptive Capacity, Learning, and Performance in International Joint Ventures," *Strategic Management Journal*, 22 (12), 1139–61.

Lee, Ruby P. and Rajdeep Grewal (2004), "Strategic Responses to New Technologies and Their Impact on Firm Performance," *Journal of Marketing*, 68 (October), 157–71.

Levitt, Theodore (1969), *The Marketing Mode*. New York: McGraw-Hill.

Li, Haiyang and Kwaku Atuahene-Gima (2001), "Product Innovation Strategy and the Performance of New Technology Ventures in China," *Academy of Management Journal*, 44 (6), 1123–34.

Li, Tiger and Roger Calantone (1998), "The Impact of Market Knowledge Competence on New Product Advantage: Conceptualization and Empirical Examination," *Journal of Marketing*, 58 (October), 13–29.

Luo, Yadong (2001), "Antecedents and Consequences of Personal Attachment in Cross-Cultural Cooperative Ventures," *Administrative Science Quarterly*, 46 (2), 177–201.

Maltz, Elliot and Ajay Kohli (1996), "Market Intelligence Dissemination Across Functional Boundaries," *Journal of Marketing Research*, 33 (February), 47–61.

——— and ——— (2000), "Reducing Marketing's Conflict with Other Functions: The Differential Effects of Integrating Mechanisms," *Journal of the Academy of Marketing Science*, 28 (4), 479–92.

Menon, Anil, Sundar G. Bharadwaj, Phani Tej Adidam, and Steven W. Edison (1999), "Antecedents and Consequences of Marketing Strategy Making: A Model and a Test," *Journal of Marketing*, 63 (April), 18–40.

——— and P Rajan Varadarajan (1992), "A Model of Marketing Knowledge Use Within Firms," *Journal of Marketing*, 56 (October), 53–71.

Moorman, Christine (1995), "Organizational Market Information Processes: Cultural Antecedents and New Product Outcomes," *Journal of Marketing Research*, 32 (August), 318–35.

——— and Roland T. Rust (1999), "The Role of Marketing," *Journal of Marketing*, 63 (Special Issue), 180–97.

———, Gerald Zaltman, and Rohit Deshpandé (1992), "Relationships Between Providers and Users of Market Research: The Dynamics of Trust Within and Between Organizations," *Journal of Marketing Research*, 29 (August), 314–29.

Morgan, Robert M. and Shelby D. Hunt (1994), "The Commitment–Trust Theory of Relationship Marketing," *Journal of Marketing*, 58 (July), 20–38.

Narver, John C. and Stanley F. Slater (1990) "The Effect of a Market Orientation on Business Profitability," *Journal of Marketing*, 54 (October), 20–35.

Nonaka, Ikujiro (1994), "A Dynamic Theory of Organizational Knowledge Creation," *Organization Science*, 5 (1), 14–37.

Nunnally, Jum C. (1978), *Psychometric Theory*, 2d ed. New York: McGraw-Hill.

Peng, Mike W. and Yadong Luo (2000), "Managerial Ties and Firm Performance in a Transition Economy: The Nature of a Micro-Macro Link," *Academy of Management Journal*, 43 (3), 486–501.

Rindfleisch, Aric and Christine Moorman (2001), "The Acquisition and Utilization of Information in New Product Alliances: A Strength-of-Ties Perspective," *Journal of Marketing*, 65 (April), 1–18.

——— and ——— (2003), "Interfirm Cooperation and Customer Orientation," *Journal of Marketing Research*, 40 (November), 421–36.

Ruekert, Robert and Orville Walker (1987), "Marketing's Interaction with Other Functional Units: A Conceptual Framework and Empirical Evidence," *Journal of Marketing*, 51 (January), 1–19.

Rust, Roland T., Tim Ambler, Gregory S. Carpenter, V. Kumar, and Rajendra K. Srivastava (2004), "Measuring Marketing Productivity: Current Knowledge and Future Directions," *Journal of Marketing*, 68 (October), 76–89.

Sinkula, James (1994), "Market Information Processing and Organizational Learning," *Journal of Marketing*, 58 (January), 35–45.

Sivadas, Eugene and F. Robert Dwyer (2000), "An Examination of Organizational Factors Influencing New Product Success in Internal and Alliance-Based Processes," *Journal of Marketing*, 64 (January), 31–49.

Slater, Stanley F. and John C. Narver (1994), "Does Competitive Environment Moderate the Market Orientation–Performance Relationship?" *Journal of Marketing*, 58 (January), 46–55.

Srivastava, Rajendra K., Tasadduq A. Shervani, and Liam Fahey (1998), "Market-Based Assets and Shareholder Value: A Framework for Analysis," *Journal of Marketing*, 62 (January), 2–18.

Szulanski, Gabriel (1996), "Exploring Internal Stickiness: Impediments to the Transfer of Best Practice Within the Firm," *Strategic Management Journal*, 17 (Winter Special Issue), 17–43.

Tsai, Wenpin (2002), "Social Structure of 'Coopetition' Within a Multiunit Organization: Coordination, Competition, and Intraorganizational Knowledge Sharing," *Organization Science*, 13 (2), 179–90.

Uzzi, Brian (1997), "Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness," *Administrative Science Quarterly*, 42 (1), 35–67.

——— (1999), "Embeddedness in the Making of Financial Capital: How Social Relations and Networks Benefit Firms Seeking Financing," *American Sociological Review*, 64 (August), 481–505.

Van Bruggen, Gerrit, Gary Lilien, and Manish Kacker (2002), "Informants in Organizational Marketing Research: Why Use Multiple Informants and How to Aggregate Responses," *Journal of Marketing Research*, 39 (November), 469–78.

Vargo, Stephen L. and Robert F. Lusch (2004), "Evolving to a New Dominant Logic for Marketing," *Journal of Marketing*, 68 (January), 1–17.

Von Hippel, Eric (1987), "Cooperation Between Rivals: Informal Know-How Trading," *Research Policy*, 16 (6), 291–302.

Walters, Peter G. and Saeed Samiee (2003), "Marketing Strategy in Emerging Markets: The Case of China," *Journal of International Marketing*, 11 (1), 97–106.

Zahra, Shaker and Gerard George (2002), "Absorptive Capacity: A Review, Reconceptualization, and Extension," *Academy of Management Review*, 27 (2), 185–203.

Zeng, Ming and Xiao-Ping Chen (2003), "Achieving Cooperation in Multiparty Alliances: A Social Dilemma Approach to Partnership Management," *Academy of Management Review*, 28 (4), 587–605.

Zhou, Kevin, Chi Kin Yim, and David Tse (2005), "The Effects of Strategic Orientations on Technology- and Market-Based Breakthrough Innovations," *Journal of Marketing*, 69 (April), 42–60.

Jennifer J. Argo, Darren W. Dahl, & Andrea C. Morales

Consumer Contamination: How Consumers React to Products Touched by Others

Although consumers like to touch products while shopping, the authors propose a theory of consumer contamination, positing that consumers evaluate products previously touched by other shoppers less favorably. The authors test the theory by manipulating cues that increase the salience that consumer contact has occurred. Furthermore, the authors investigate the role of disgust as the underlying mechanism of the theory.

A recent article in *The New York Sun* (Vanech 2005) highlights the difficulty that cosmetic retailers have in managing consumers' touching and interacting with their product offerings. Matthew Waitesmith, vice president of MAC cosmetics, is quoted as saying "we never double-dip" (p. 20), in reference to stringent policies the company uses in an effort to alleviate consumer concern over trial and sampling of its retail products. Similarly, in the best-selling book, *Why We Buy*, Paco Underhill (2000) discusses the challenges that retailers face in effectively managing the impact of consumers' touching display products. He notes that consumers often open the packaging of products to touch and feel them but usually choose to buy products that have been or appear to be untouched by others. Indeed, it is common for consumers to try on an article of clothing or flip through a magazine but, after deciding to make a purchase, select a "fresh" one from the back of the display. It is clear that there is a strong consumer aversion to touched products, and it is difficult to manage consumer touching behavior. However, little research has investigated the reason for this phenomenon or has provided managers with insight into how to address its implications.

Recently, marketing researchers have found that information obtained through physical touch plays a central role in the evaluation of consumer products (Mooy and Robben 2002; Peck and Childers 2003). The importance of physical touch in consumption has been cited as one of the critical factors that has limited the adoption of online shopping, an environment that is void of tactile information (Alba et al. 1997; Citrin et al. 2003; McCabe and Nowlis 2003). Although prior research suggests that touch has a positive impact on consumer information search and product evaluations, the goal of the current research is to show that consumer contact with products may actually be a double-edged sword for marketers. Specifically, we develop a theory of consumer contamination, which we define as contamination from consumer touching. In the product context we investigated herein, this behavior results in negative implications for a touched product. Importantly, in our research, we test the impact of consumer contamination by examining cases in which products are unharmed objectively by contact with other consumers, but nevertheless the contact is still viewed as a contamination event. This differs from actual physical contamination in which a product becomes soiled or dirty as a result of consumer contact.

Using the law of contagion (Frazer [1890] 1959; Mauss [1902] 1972; Tylor [1871] 1974) as a theoretical framework, we propose that when consumers become cognizant that another consumer has previously touched a product, their evaluations of and purchase intentions for the product decrease. We test this prediction in a retail shopping environment by manipulating factors that increase the salience that consumer contact has occurred. We refer to such factors as "contamination cues." In particular, our research investigates the effects of three contamination cues—proximity to contact, time elapsed since contact, and the number of contact sources—on consumers' product evaluations and purchase intentions. In addition to establishing the impact of consumer contamination in a retail context, we also provide insight into the underlying process through which consumer contamination operates by investigating the role of disgust as the driving mechanism. Our research reveals that the increased salience of contact through a variety of contamination cues negatively affects consumer evaluations of touched products.

Conceptual Background

Law of Contagion

In the late 1800s and early 1900s, anthropologists developed a conceptual framework for the laws of sympathetic

---

Jennifer J. Argo is Assistant Professor of Marketing, School of Business, University of Alberta (e-mail: jennifer.argo@ualberta.ca). Darren W. Dahl is Associate Professor of Marketing, Sauder School of Business, University of British Columbia, Vancouver (e-mail: darren.dahl@sauder.ubc.ca). Andrea C. Morales is Assistant Professor of Marketing, W.P. Carey School of Business, Arizona State University (e-mail: acmorales@asu.edu). The authors contributed equally to this research, and the order of authorship is alphabetical. The authors thank Gavan Fitzsimons and Gal Zauberman for their helpful comments. The authors gratefully acknowledge financial support from the Social Sciences and Humanities Research Council of Canada and the Southam/Edmonton Journal Fellowship. The authors also thank the University of Alberta Bookstore for its assistance in conducting the experiments.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

magic (Frazer [1890] 1959; Mauss [1902] 1972; Tylor [1871] 1974). These laws summarize a pattern of beliefs and practices that exist over several traditional cultures to describe how the world works. One of the central laws of sympathetic magic is the law of contagion. According to this law, when a source (person/object) and a recipient (another person/object) come into direct or indirect contact, the source influences the recipient (Rozin and Nemeroff 1990). This influence arises because when the two touch, the source transfers its essence (i.e., some or all of its essential properties) to the recipient (Nemeroff and Rozin 1994). Notably, the essence does not need to be visible to the naked eye and can be composed of invisible entities (Rozin and Nemeroff 1990). According to the law, the source's essence remains a part of the recipient even after contact has been broken (Rozin and Nemeroff 1990).

Previous research has indicated that the law of contagion can play a significant role in an interpersonal context. For example, in anthropological research on the Hua tribe of Papua New Guinea, Meigs (1984) finds that when a Hua woman marries and moves to her husband's community, no one in the community will consume the food she produces and/or serves. This refusal arises because the villagers believe that the woman's essence has been transferred to the food; if they accept the food, they will become polluted with her properties (Meigs 1984).

In a different context, correlational research has shown that people can have both positive and negative responses to objects that have been touched by others. Although people are reluctant and uncomfortable coming into physical contact with an object that has been used by a disliked person (e.g., a laundered sweater that belonged to Adolf Hitler), this research shows that they sometimes evaluate objects used by people they like more favorably (e.g., a hamburger with a bite taken out of it by their lover; Nemeroff 1995; Rozin, Markwith, and McCauley 1994; Rozin, Millman, and Nemeroff 1986; Rozin et al. 1989). However, the processes that are believed to underlie such negative and positive effects from contact are not identical. Negative contamination (i.e., the value of an object is decreased through contact) has been shown to be the result primarily of a residue model that results from physical contact. Conversely, positive contagion (i.e., the value of an object is increased through contact) has been found to be the result strictly of a symbolic interaction model that is based on interpersonal/moral factors (Nemeroff and Rozin 1994). Thus, previous research shows that negative contamination effects are driven primarily by a transfer of essence, as in the Hua tribe example, and that positive contamination effects arise primarily because people want to be associated with an item that belongs to someone about whom they have strong, positive feelings (Nemeroff and Rozin 1994). Although both positive and negative contamination are possible in the retail context under investigation in this article, we focus specifically on negative contamination effects.

Negative Contamination in a Retail Context

In O'Reilly and colleagues' (1984) study, 76% of survey respondents indicated that they would not buy used underclothing, and 20% refused to purchase used overcoats. The justification for the low desirability of the used products was related to a fear of contamination. Because the original owners of the used products remained anonymous to consumers, these feelings could not be based on associations but instead suggest a belief that the clothing was contaminated simply because it had been handled and owned by someone else previously. Indeed, there are many cases in which contact with consumers actually causes products to become less desirable to other consumers as a result of physical contamination from incidental touching or usage (e.g., floor-room models, demonstration products). In the current research, we argue that consumers generalize this notion for even brief and incidental contact that does not involve formal usage, ownership, or actual physical contamination.

This generalized contamination effect has negative implications for products that have been handled by consumers. Much like second-hand items or demonstration products that suffer devaluation (Kotler and Mantrala 1985), contamination that arises from incidental consumer contact should also lead consumers to devalue touched products. This devaluation should be manifested by lower evaluations of and purchase intentions for touched products than for untouched products.

Note also that consumer contact with a product does not need to be formally observed to produce contamination effects (Rozin, Markwith, and McCauley 1994; Rozin et al. 1989). Rarely does a person actually see who has put clothes on consignment or who has previously test-driven an automobile. In these cases, consumers infer that contact has occurred. Thus, simply a cue in the retail environment that signals to a consumer that someone else has previously touched the product is necessary for consumer contamination to occur. Given this, we expect that perceived contact between consumers and products is sufficient, and observed physical contact is not essential for the elicitation of consumer contamination effects. More formally,

> P_1: Perceived contact between a consumer and a product results in other consumers having less favorable evaluations of the touched product.

Implicit to our proposition is the presence of environmental signals that provide cues to the consumer that a product has been touched. These signals, or contamination cues, can be relayed through a variety of sources, including, but not limited to, other consumers, sales staff, merchandising placement, and product characteristics. These cues increase the salience of the contamination and thus increase the likelihood of consumer contamination. Angyal (1941) provides support for this idea by noting that the greater the salience or intimacy of a contamination event, the greater is the impact of contamination effects. Therefore, the more clearly a contamination cue signals that contact between a consumer and a product has taken place, the greater is the effect of contamination on other consumers' evaluations of that product. Moreover, the greater the intimacy of contact, the greater is the likelihood of contamination occurring. This implies that contamination effects differ depending on the type of cue present and the type of product, because the degree of intimacy changes by product category. In the current research, however, we focus only on the effects of different contamination cues.

Although there are likely to be several possible contamination cues present in the retail environment, we focus our empirical investigation on three specific cues: proximity to contact, time elapsed since contact, and number of contact sources. In the next section, we discuss and develop specific hypotheses that are related to these contamination cues. Furthermore, drawing on previous work on contamination (Rozin and Fallon 1987; Rozin, Millman, and Nemeroff 1986), we also forward a prediction about the role of disgust as a central mechanism in the incidence of consumer contamination in a retail context.

Hypotheses

Proximity to Contact as a Contamination Cue

The first contamination cue we examine is proximity to contact, which we define as the physical location of the product relative to where it would be touched by other consumers in a retail setting. To understand how proximity to contact influences evaluations and purchase intentions of touched products, we draw from social impact theory (SIT; Latané 1981), which proposes that the presence or action of a real, implied, or imagined social presence affects a target. This impact results from three forces: immediacy (i.e., proximity), number (i.e., how many), and strength (i.e., importance; Latané 1981). One of the theory's principles predicts that the impact of each of the forces increases as a power function. Thus, in terms of proximity, the greatest impact arises from close proximity as opposed to being farther away (Argo, Dahl, and Manchanda 2005; Latané and Wolf 1981). Extending this principle to consumer contamination, we predict that contamination salience will be greater the closer a product is situated to a location of contact. We expect that this will happen because as the product moves closer in proximity to a contact location, consumers will perceive that there is an increase in the probability that the product has been touched. This expectation is consistent with observations from retail settings in which consumers often select products, such as magazines, that are located at the bottom or back of a display (i.e., where there is a lower probability that the product has been touched because it is located farther out of a consumer's reach). Thus, we expect that as contamination salience increases through proximity to contact, there will be an increase in the impact of consumer contamination on product evaluations and purchase intentions. On the basis of this discussion, we expect that this impact will be negative in valence.

H_1: Consumer product evaluations and purchase intentions will be lower if a product is physically proximate (versus not physically proximate) to the location where another consumer is believed to have had contact with it.

Time Elapsed Since Contact as a Contamination Cue

The second contamination cue we consider is the amount of time that has elapsed since consumer contact has occurred. Unlike proximity to contact, it is unclear how time influences product evaluations and purchase intentions. The law of contagion suggests that when contact has been made between a source and a recipient, the source continues to influence the recipient indefinitely—"once in contact, always in contact" (Frazer [1890] 1959; Rozin, Markwith, and McCauley 1994). In other words, after an object has been contaminated, it should remain that way, even after the two items are no longer touching (Frazer [1890] 1959). Preliminary evidence for this idea has been found in two correlational studies. In the first study, Rozin and colleagues (1989) find that attempts to cleanse a contaminated item by washing it failed to eliminate the effects of contamination completely. In the second study, Hejmadi, Rozin, and Siegal (2004) demonstrate that children believed that juice continued to be contaminated even after a contaminant had been removed and the juice was purified. These findings provide some suggestive evidence that consumers will believe that a product is contaminated by physical contact with another consumer, regardless of the amount of time that has elapsed since it was last touched. Thus, the impact of consumer contamination on product evaluations and purchase intentions should be consistent across time.

However, recent research on construal-level theory suggests that consumers' responses to a contamination event vary at different points in time (Liberman and Trope 1998; Trope and Liberman 2003). Trope and Liberman (2003) examine how temporal distance systematically alters the way people mentally construe future events. They posit that when people consider distant-future events, they form high-level construals (i.e., abstract representations). In contrast, when people consider near-future events, they form low-level construals that involve more concrete and vivid details. Although Trope and Lieberman focus exclusively on the impact of temporal distance for future events, they suggest that construal theory principles should apply to past events as well. Therefore, when people consider distant-past events, similar to the distant-future events, they will form higher-level, less concrete construals about the event than when they consider near-past events. Given this, it would be expected that the salience of consumer contamination would be mitigated as time passes. In other words, the shorter the time elapsed since consumer contact, the stronger is the consumer contamination effects realized. We offer two competing hypotheses:

H_{2a}: Consumer product evaluations and purchase intentions will be lower if a product is believed to have had contact with another consumer, regardless of the time elapsed since contact.

H_{2b}: Consumer product evaluations and purchase intentions will be lower if a product is believed to have had contact with another consumer; this effect will be mitigated as time elapses since contact.

Number of Contact Sources as a Contamination Cue

The third contamination cue we examine is the number of contact sources. Specifically, we consider how consumer contamination effects are influenced by the number of consumers believed to have had physical contact with a product. As we mentioned previously, SIT predicts that the impact of the number of people who constitute a real, imagined, or implied social presence will increase as a power function. Therefore, we expect that an increase in the number of sources that a consumer believes have previously touched a product will increase the contamination effects.

Although SIT predicts that contamination effects should be stronger (i.e., lower product evaluations) when more consumers are believed to have come in contact with a product, multiple contact sources might actually increase product evaluations by signaling that the product is popular. Thus, rather than lower the value of the product, contact by multiple consumers may heighten evaluations and purchase intentions of a touched product by emphasizing its popularity. We hypothesize the following competing predictions:

H_{3a}: Consumer product evaluations and purchase intentions will be lower if a product is believed to have had contact with many (versus one) consumers.

H_{3b}: Consumer product evaluations and purchase intentions will be higher if a product is believed to have had contact with many (versus one) consumers.

Feelings of Disgust as the Underlying Mechanism

The previous discussion argues that as contamination becomes more salient to consumers, they are more likely to experience consumer contamination effects. What is unclear, however, is the underlying mechanism that drives consumers to respond unfavorably to products that have been touched by other consumers. Previous research has suggested that the law of contagion operates in the Western world in the domain of disgust (Rozin, Millman, and Nemeroff 1986), which is defined as revulsion in response to an offensive object (Rozin and Fallon 1987). Feelings of disgust have been shown to underlie contamination effects in various contexts, including food rejection, disease, and moral taint (Rozin, Haidt, and McCauley 1993; Rozin, Markwith, and Nemeroff 1992). Furthermore, in general, disgust manifested in rejection and/or lower evaluations of a contaminated object (Rozin and Fallon 1987; Rozin, Haidt, and McCauley 1993; Rozin and Nemeroff 1990; Rozin, Millman, and Nemeroff 1986).

Given these findings, we expect that consumers who perceive that physical contact has occurred between another consumer and a product will experience feelings of disgust toward the contaminated product. In turn, these feelings of disgust will translate into negative consumer evaluations and purchase intentions. Thus, disgust is the primary driving force that motivates a change in consumer evaluations for products that are touched by others.

H_4: Disgust mediates the influence of consumer contamination on product evaluations and purchase intentions.

We test these hypotheses in three field-study experiments. In the first study, we investigate the role of the proximity to contact (H_1) in facilitating contamination effects. Study 2 broadens the investigation by adding a manipulation of time elapsed since contact and provides an explicit test of the competing hypotheses put forth in H_2. Finally, Study 3 examines how the number of contact sources (H_3) influences consumer contamination. In each experiment, we measure feelings of disgust as an underlying motivator (H_4) in producing the contamination effects that consumers experience.

Study 1

Overview

The purpose of this study is to establish that consumer contamination operates in a retail setting, negatively influencing evaluations of products that have been touched by other consumers. Furthermore, we empirically test H_1 and H_4. Study 1 consists of a one-way between-subjects experimental design with proximity to contact (close versus medium versus far) as the experimental factor. We randomly assigned 70 undergraduate students from a large North American university to one of the three conditions; they received course credit for their participation.

Method

Procedure. Participants were tested individually at the university student center. When they arrived, they were informed that the objective of the study was to gather marketing research information for the university bookstore. Participants were told to visit the store, look around to get an impression of the store (three to five minutes), and then complete an assigned task. To identify the task they needed to complete while in the store, they were instructed to select an envelope that would identify a task for them. They were told that the tasks ranged from making a purchase to trying or using a product. Unbeknownst to the participants, each envelope indicated that they would be required to perform the same task: to try on a specific T-shirt located in the store. The envelope included a picture of the target T-shirt and instructed participants to find a sales associate in the clothing section who would help locate the proper article of clothing. After selecting the envelope, participants proceeded to the bookstore. At this time, the experimenter text-messaged a trained confederate located in the store to indicate that the participant was en route and to provide a description of him or her. The confederate assumed the role of a sales associate at the bookstore and was responsible for implementing the proximity-to-contact manipulation. One male and one female confederate alternated the role, and both were blind to the experimental hypotheses.[1]

In all conditions, after participants located the sales associate in the store and showed him or her the photograph of a one-of-a-kind T-shirt, the sales associate said, "We only have one left in stock." In the close condition, the sales associate finished the sentence with "and someone is trying it on." The sales associate then took the participant over to the dressing room to wait, and a couple of seconds later, a second confederate, who assumed the role of another shopper (same gender as the participant), exited the dressing room, leaving the T-shirt in the room. In the medium condition, the sales associate completed the sentence with "and it is over here on the return rack" and then took the participant to a return rack located near the dressing room. In the far condition, the sales associate finished the sentence with "and it is just over here on the rack" and took the participant to a regular display rack located a few feet away from the return rack. Note that at no time in any of the conditions did either confederate actually touch the target T-shirt. Physical contact with the target T-shirt was implied only by the proximity-to-contact manipulation.

---

[1] The sales associate was also responsible for monitoring the cleanliness of the T-shirt and replaced it with a new one after every second participant or if it became soiled.

After finishing the task and returning to the experimenter, participants completed a questionnaire. Participants were then debriefed and were thanked for their participation.

Dependent variables. The questionnaire contained both the dependent variables of interest and questions consistent with the cover story. We measured overall product evaluations using five seven-point scales. The items included "bad/good," "negative/positive," "undesirable/desirable," "unfavorable/favorable," and "dislike/like." We averaged the responses to form an evaluation index ($\alpha = .95$). We measured purchase intentions using a seven-point scale that asked participants to indicate how likely they would be to buy the product (1 = "very unlikely to buy," 7 = "very likely to buy").

We also asked participants to indicate how they felt during the shopping experience using a battery of both positive and negative affect terms (1 = "not at all," 7 = "very"). Embedded in this list were four items intended to measure feelings of disgust (Haidt, McCauley, and Rozin 1994; Smith and Ellsworth 1985) and other negative emotional responses that could be elicited by the shopping experience and could potentially account for changes in product evaluations. We created a disgust index by averaging the following four items: disgusted, revolted, unclean, and gross ($\alpha = .76$). To rule out the possibility that negative emotions other than disgust accounted for the results, we created another index for other negative emotions by averaging the following five items: frustrated, bad, annoyed, angry, and mad ($\alpha = .85$).

To determine whether participants' reactions toward the product were a direct result of specific physical contamination or a more general contamination effect, we asked them, "How dirty was the product you used/tried?" We averaged two seven-point scales ("dirty, unsanitary" and "not at all/very") to create a dirty index ($r = .70$, $p < .01$). After completing these measures, participants completed a manipulation check for proximity to contact (location recall measure) by answering an open-ended question that asked where they found their target product. Participants also indicated their gender, age, and major and completed an open-ended suspicion probe. Responses to the demographic measures (including match/mismatch on gender relative to the confederate) were not significant in this study or any of the subsequent studies; thus, we do not discuss this issue further.

Results

Preliminary analyses. Of the 70 participants who took part in the study, 67 successfully completed the assigned task. The sample used in the subsequent analysis consisted of 67 participants (males = 30, females = 37); cell sizes ranged from 21 to 24. An analysis of the manipulation check for proximity indicated that all participants were cognizant of where the product was located. An examination of responses to the suspicion probe revealed that no participants were aware of the experimental hypotheses.

Test of hypotheses. H_1 predicted that product evaluations and purchase intentions would decrease as the product became closer to where contact with another consumer took place (i.e., contamination salience increased). An analysis of variance (ANOVA) using proximity to contact as the independent variable and the evaluation index as the dependent variable produced a significant main effect ($F(2, 64) = 22.95$, $p < .001$). Post hoc tests indicated that reducing proximity to contact results in a linear decrease in product evaluations. Participants evaluated the T-shirt significantly more positively when it was located on the regular rack ($M_{far} = 5.20$) than when it was on the return rack ($M_{medium} = 3.86$, $t(64) = 3.24$, $p < .01$) or in the dressing room ($M_{close} = 2.47$, $t(64) = 6.77$, $p < .001$), and they evaluated it significantly more positively when it was located on the return rack than when it was in the dressing room ($t(64) = 3.50$, $p < .001$).

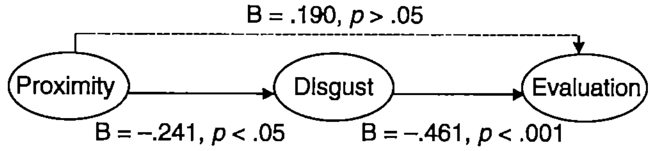
To determine whether proximity to contact also significantly influenced participants' purchase intentions, we conducted a second ANOVA. Again, we found a significant effect ($F(2, 64) = 9.04$, $p < .001$). Post hoc comparisons showed that participants' purchase intentions were lower as the proximity to the contact location became closer ($M_{close} = 2.46$; $M_{medium} = 3.64$, $t(64) = 2.20$, $p < .05$; close versus far: $M_{far} = 4.76$, $t(64) = 4.25$, $p < .001$; medium versus far: $t(64) = 2.03$, $p < .05$). Thus, we found support for H_1.

To determine whether feelings of disgust mediated the impact of proximity to contact on product evaluations and purchase intentions (H_4), we conducted a mediation analysis. Following the work of Baron and Kenny (1986), we found that proximity to contact significantly predicted feelings of disgust ($F(2, 64) = 28.02$, $p < .001$; $M_{close} = 4.11$, $M_{medium} = 2.35$, $M_{far} = 1.27$). Next, we estimated a path analysis that included the direct path for the effect of proximity to contact on product evaluations and an indirect path mediated by disgust. The results produced a significant indirect path, indicating that disgust fully mediated the effect of the independent variable on product evaluations (see Figure 1, Panel A). We also conducted a path analysis to establish whether disgust significantly mediated purchase intentions. A second mediation analysis indicated that disgust fully mediated the effect of proximity to contact on purchase intentions (see Figure 1, Panel B). We conducted an additional path analysis to determine whether product evaluations mediated the impact of proximity to contact on purchase intentions. As we show in Figure 1, Panel C, product evaluations fully mediated the relationship.

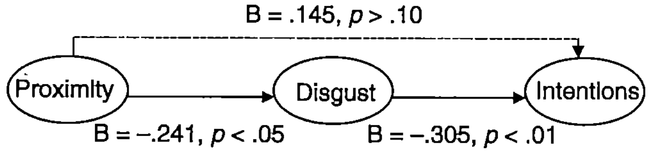
Additional analyses. To rule out the possibility that contamination made participants experience other negative emotions and that these negative emotions caused the lower evaluations of the contaminated products, we conducted an analysis using the negative emotion index. The ANOVA indicated that proximity to contact failed to significantly

FIGURE 1
Study 1: Path Analysis

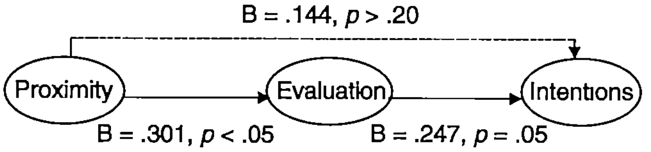
A: Mediating Role of Disgust on Evaluations



B: Mediating Role of Disgust on Intentions



C: Mediating Role of Evaluations on Intentions



influence other negative emotions (F(2, 64) = 1.71, $p > .10$), providing evidence that feelings of disgust are specific in the facilitation of consumer contamination effects. A more general, negative reaction does not appear to drive the lower evaluations and purchase intentions for a product that has been touched by other consumers.

We also conducted an analysis using the dirty index as a dependent variable (F(2, 64) = 1.89, $p > .10$). Null effects for this measure indicate that consumers do not perceive actual physical contamination from the consumer touching that is inherent in our manipulation but rather experience a more general contamination effect that directly affects their evaluations and purchase intentions for the product. An analysis in subsequent studies (i.e., Studies 2 and 3) also shows null effects for both the negative emotion and dirty indexes; thus, we do not discuss these measures further.

Discussion of Study 1

Study 1 provides the first empirical evidence in support of the existence of consumer contamination. The results show that when consumers receive a signal that another consumer has touched a product, they lower both their evaluations and their purchase intentions for that product. Notably, participants never actually observed other people touching the product; instead, they inferred contact through the contamination cue of proximity to contact. By changing where the product was located in the store, we also manipulated the salience of contamination. In particular, when the product was located closer to the perceived point of contact, participants' evaluations of and purchase intentions for the product were lower. This decrease in evaluations and intentions appears to be driven by feelings of disgust that arise when another shopper is believed to have contaminated a product through contact.

Although the results of Study 1 are consistent with a theory of consumer contamination, it is possible that the lower product evaluations and purchase intentions in the dressing room and return rack (compared with the regular rack) conditions were not the outcome of contamination but rather the result of consumers thinking that the product they were asked to try on had been previously rejected. In other words, the location of the product (i.e., either in the dressing room or on the return rack) may have signaled to participants that the product was deficient because someone else had previously tried it on but decided not to purchase it. To test this alternative explanation, we implemented an additional experimental condition (rejection condition).

In the rejection condition, as in the far condition, the sales associate indicated that the product was located on a regular rack in the store. However, when the sales associate and the participant arrived at the rack, they found that it was empty. After the sales associate vocalized that it appeared as though someone else already had the T-shirt, another shopper (i.e., a confederate), who was standing nearby holding the target T-shirt and a couple other items by the hangers, looked over and said, "I was going to try it on, but I don't want it after all." Without ever making contact with the shirt, he or she then handed the participant the one-of-a-kind T-shirt on the hanger. Independent sample t-tests compared the rejection condition and the far condition, and the results indicated that the two conditions did not differ significantly for either product evaluations (t(38) = .74, $p > .20$, $M_{rejection} = 4.87$) or purchase intentions (t(38) = .35, $p > .20$, $M_{rejection} = 4.58$). The absence of statistical significance suggests that negative product evaluations and purchase intentions in the dressing room and return rack conditions are consistent with consumer contamination and are not explained by another consumer rejecting the product.

Another possible explanation for the results of Study 1 that cannot be addressed with the current findings is that the location of the product not only manipulated proximity to contact but also may have changed perceptions about the amount of time elapsed since contact. Specifically, when the product was located in the dressing room or on the return rack, consumers may have thought that the product was more recently contaminated than when it was located on the regular rack. In turn, knowledge about the time elapsed since contact may have subsequently affected product evaluations and purchase intentions. We tease apart these two competing explanations in Study 2 by adding an explicit manipulation of time elapsed since contact from another consumer.

Study 2

Overview

We conducted Study 2 to investigate the role of time elapsed since contact in the realization of consumer contamination. This study empirically tests the competing hypotheses H_{2a} and H_{2b}. Furthermore, in this study, to clarify the role of proximity to contact and time elapsed since contact in producing contamination effects, we manipulate both contamination cues separately. Participants in the study included 95 students from a large North American university. We randomly assigned them to a condition in a 2 (proximity to contact: medium versus far) × 2 (time elapsed since contact: short versus long) + 1 (control: no touch) between-subjects experimental design. The 48 male and 47 female participants received $10 for compensation for their participation. Cell sizes ranged from 18 to 20.

Method

Procedure. We used the same procedure as that in Study 1, but with two exceptions: First, we used only two levels of proximity to contact, medium and far. We manipulated both the medium and the far conditions as in Study 1. Second, to manipulate time elapsed since contact, the sales associate indicated the amount of time that had passed since the T-shirt was last touched by another consumer. In the short condition, the sales associate told the participant that the T-shirt had just been tried on, whereas in the long condition, the sales associate indicated that the T-shirt had not been tried on for a few days. We included an additional group in the design to serve as a control group. Participants in this condition found the target product located on the regular rack (similar to the far proximity-to-contact condition) but were not provided with any contact information. The manipulation check for time elapsed since contact used two nine-point scales that asked, "How long do you think the product was in the same exact location where you found it?" (1 = "very short period of time," 9 = "very long period of time") and "How much time do you think the product was in the same exact location where you found it?" (1 = "minutes," 9 = "months"). We averaged these two items and combined them to form a time index ($r = .73$, $p < .001$). Dependent measures followed those in Study 1 (evaluation index $\alpha = .96$, disgust index $\sigma = .97$, location recall measure, and purchase intentions).

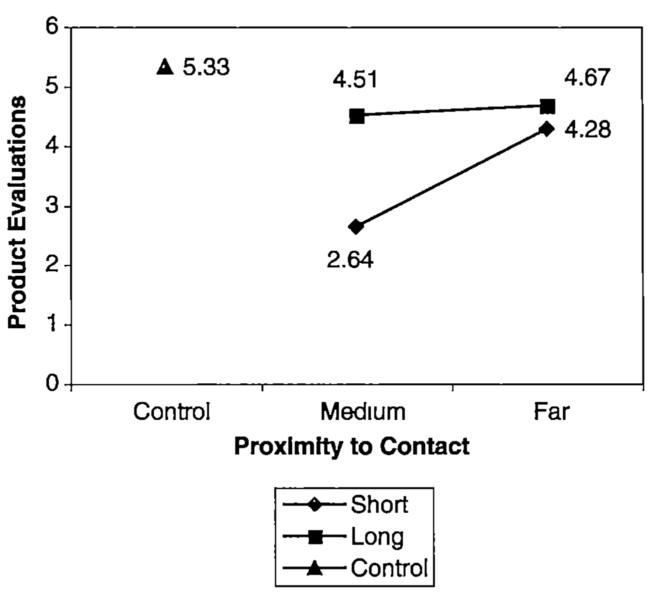
Results

Preliminary analyses. All participants were cognizant of the product's location. A 2 × 2 ANOVA with the time index as the dependent variable produced only a significant main effect for time elapsed since contact ($F(1, 71) = 13.25$, $p = .001$; $M_{long} = 5.26$, $M_{short} = 3.63$).[2] Thus, the manipulation checks were successful. We also examined responses to the suspicion probe, which revealed that participants were unaware of the experimental hypotheses.

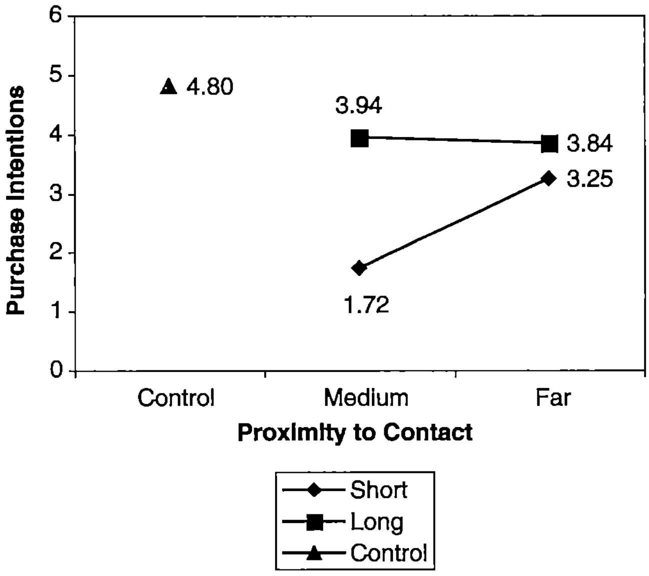
Test of hypotheses. Recall that H_2 put forth competing predictions about the role of time elapsed since contact for consumer contamination. H_{2a} proposed that the passing of time would not influence contamination, whereas H_{2b} proposed that the passage of time would mitigate contamination effects. A 2 (proximity to contact: medium versus far) × 2 (time elapsed since contact: short versus long) ANOVA with the evaluation index as the dependent variable produced a main effect for proximity to contact ($F(1, 71) = 8.09$, $p < .01$; $M_{medium} = 3.58$, $M_{far} = 4.47$) and a main effect for time elapsed since contact ($F(1, 71) = 12.78$, $p = .001$; $M_{long} = 4.60$, $M_{short} = 3.51$). However, these main effects were qualified by a significant interaction between the two factors ($F(1, 71) = 5.43$, $p < .05$; $M_{medium\ and\ short} = 2.64$, $M_{medium\ and\ long} = 4.51$, $M_{far\ and\ short} = 4.28$, $M_{far\ and\ long} = 4.67$; see Figure 2, Panel

FIGURE 2
Study 2: Influence of Proximity to Contact and Time Elapsed Since Contact

A: Product Evaluations Index Scores



B: Purchase Intentions Score



---

[2] We conducted a one-way ANOVA with the control group and the amount of time conditions for the time delay index ($F(2, 92) = 15.95$, $p < .001$; $M_{control} = 6.14$; control versus long: $t(92) = 2.24$, $p < .05$; control versus short: $t(92) = 5.39$, $p < .001$).

A). Planned contrasts indicated that participants' evaluations of the product were significantly lower when the product had just been touched and was located on the return rack than when it was on the regular rack (t(71) = 3.68, $p < .001$). Furthermore, evaluations were significantly lower when the product was located on the return rack and had just been touched than when it was touched days before (t(71) = 4.10, $p < .001$). No other comparisons were significant (ts < 1). These findings support H_{2b} and suggest that the impact of consumer contamination disperses over time. It appears that for consumer contamination, "once in contact" does not mean "always in contact."

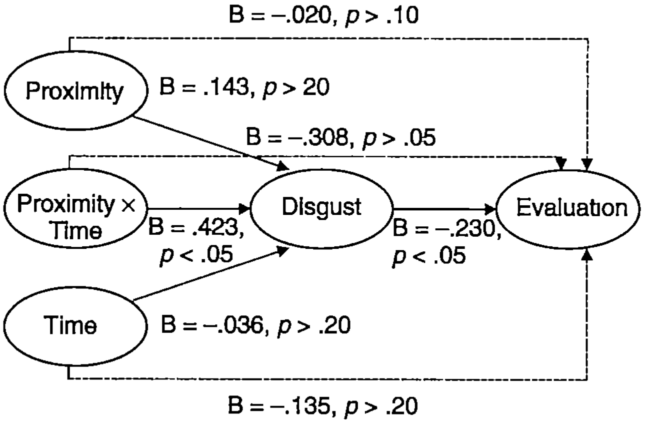
A second ANOVA indicated that purchase intentions were also significantly influenced by time elapsed since contact (F(1, 71) = 12.25, $p = .001$; $M_{long} = 3.89$, $M_{short} = 2.53$) and marginally influenced by proximity to contact (F(1, 71) = 3.14, $p < .10$, $M_{medium} = 2.83$, $M_{far} = 3.54$). However, again, these main effects were qualified by a significant interaction between proximity to contact and time elapsed since contact (F(1, 71) = 4.11, $p < .05$; $M_{medium\ and\ short} = 1.72$, $M_{medium\ and\ long} = 3.94$, $M_{far\ and\ short} = 3.25$, $M_{far\ and\ long} = 3.84$; see Figure 2, Panel B). Planned contrasts produced the same pattern as that for product evaluations. Specifically, participants were less willing to purchase the product when it had recently been touched and was located on the return rack than when the product was on the regular rack (t(71) = 2.70, $p < .01$). Furthermore, participants were less willing to purchase the product when it was on the return rack and had just been touched than when it had been days since the product had been touched (t(71) = 3.83, $p < .001$). No other contrasts were significant. Again, this result provides support for the notion that consumer contamination is not permanent but dissipates over time (H_{2b}).

A two-way ANOVA with the disgust index as the dependent variable produced two main effects (proximity to contact: F(1, 71) = 14.48, $p = .001$; $M_{medium} = 2.69$, $M_{far} = 1.47$; time elapsed since contact: F(1, 71) = 4.23, $p < .05$; $M_{long} = 1.74$, $M_{short} = 2.36$) and a significant interaction (F(1, 71) = 5.82, $p < .05$; $M_{medium\ and\ short} = 3.40$, $M_{medium\ and\ long} = 1.97$, $M_{far\ and\ short} = 1.41$, $M_{far\ and\ long} = 1.53$). To establish whether disgust mediated the impact of contamination cues on product evaluations, we conducted a mediation analysis. We estimated a path analysis that included direct paths for proximity to contact, time elapsed since contact, and their interaction on product evaluations and an indirect path that was mediated by disgust. The results produced a significant indirect path, indicating that disgust fully mediated the effects of the independent variables on product evaluations (see Figure 3, Panel A). We also conducted a path analysis for purchase intentions. Similar to product evaluations, the results revealed a significant indirect path, suggesting that disgust also fully mediated the impact of the contamination cues on intentions (Figure 3, Panel B). Finally, we conducted a path analysis to establish whether product evaluations mediated the impact of contamination cues on purchase intentions; as we indicate in Figure 3, Panel C, full mediation was present.

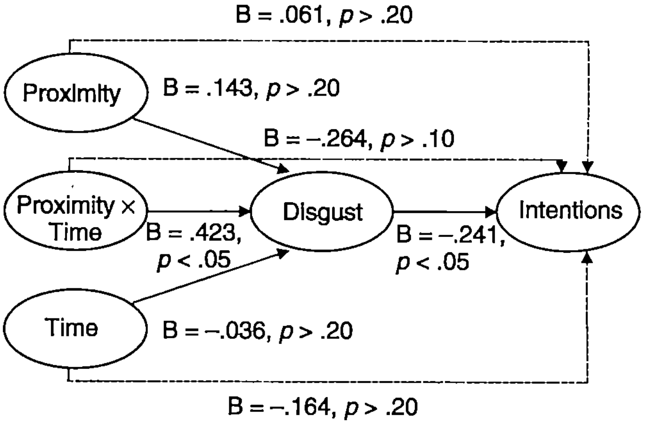
Control group. We compared the control group with the time-elapsed-since-contact conditions in a series of one-way ANOVAs, which produced significant results (product evaluations: F(2, 92) = 11.87, $p < .001$; $M_{control} = 5.33$; purchase intentions: F(2, 92) = 12.55, $p < .001$; $M_{control} = 4.80$; disgust: F(2, 92) = 4.51, $p < .05$; $M_{control} = 1.25$). Post hoc

FIGURE 3
Study 2: Path Analysis

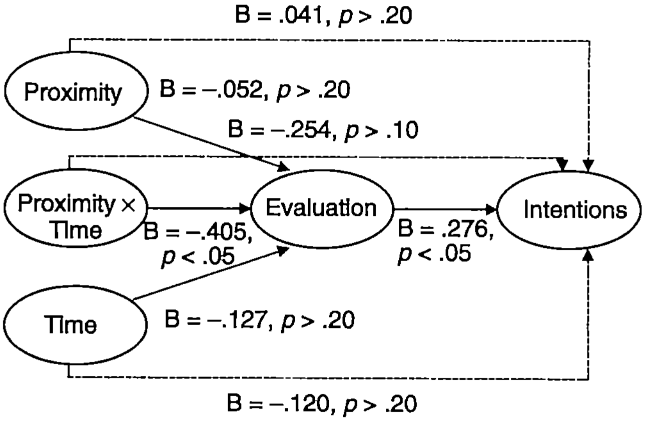
A: Mediating Role of Disgust on Evaluations



B: Mediating Role of Disgust on Intentions



C: Mediating Role of Evaluations on Intentions



tests indicated that though the control group differed significantly from the short condition (product evaluations: $t(92) = 4.63, p < .001$; purchase intentions: $t(92) = 4.75, p < .001$; disgust: $t(92) = 2.90, p < .01$), it did not differ significantly from the long condition (product evaluations: $t(92) = 1.86, p > .05$; purchase intentions: $t(92) = 1.89, p > .05$; disgust: $t(92) = 1.29, p > .10$). This provides further support for the findings that over time, the impact of contamination dissipates.

Discussion of Study 2

The findings of Study 2 provide additional support for the existence of consumer contamination in a retail setting. Consistent with Study 1, we found that as the salience of contamination increased with proximity to contact, consumer evaluations of and purchase intentions for a contaminated product decreased. However, the introduction of a second contamination cue, time elapsed since contact, moderated this finding. In particular, the results for proximity to contact only held when participants believed that other consumers had recently contaminated the product. We address the concerns identified in Study 1 by separately manipulating time elapsed since contact and defining its independent influence on consumer contamination. Following construal-level theory (Liberman and Trope 1998; Trope and Liberman 2003), when the time elapsed since contact increased, the impact of proximity to contact was mitigated. We add to the law of contagion by identifying a contamination context in which the effects of contamination appear to dissipate over time. Indeed, contamination effects do not appear to be permanent in an interpersonal, retail context.

Study 3

Overview

Studies 1 and 2 demonstrate how two contamination cues, proximity to contact and time elapsed since contact, influence consumers' evaluations of products that have been in contact with other consumers. The primary goal of Study 3 is to investigate the influence of a third contamination cue: the number of contact sources. Study 3 also extends the impact of consumer contamination to another dependent measure: willingness to pay (WTP) for the product. We randomly assigned 104 undergraduate participants to a condition in a 2 (proximity to contact: medium versus far) × 2 (number of contact sources: one versus many) + 1 (control: no touch) between-subjects experimental design. Participants received $10 as compensation for their participation.

Method

Procedure. We used the same procedure as that in Study 2, but with two exceptions: First, to manipulate the number of contact sources, in the one-contact-source condition, the sales associate informed participants that one other person had tried on the shirt, whereas in the many-contact-sources condition, the sales associate indicated that a lot of other people had tried on the shirt. We assessed the effectiveness of this manipulation by asking participants to indicate on an eight-point scale approximately how many other people had tried on/used the exact same product as them (1 = "no one," 8 = "more than seven"). Second, we introduced a new dependent measure in the study; we asked participants what they would be willing to pay for the product (i.e., dollar amount). Other measures followed those used in Study 1 (evaluation index $\alpha = .97$, disgust index $\sigma = .86$, and location recall measure).

Results

Preliminary analyses. Of the 104 participants who took part in the study, 3 were unable to complete the assigned task successfully. Thus, the sample used in the analyses reported here consisted of 101 participants (31 males and 70 females); cell sizes ranged from 18 to 22.

Analysis of the location recall question showed that all participants were cognizant of the location of the product. A 2 × 2 ANOVA with the rating of how many people had tried on the product as the dependent variable produced a significant main effect for number of contact sources ($F(1, 77) = 11.51, p < .001$; $M_{one} = 3.60$, $M_{many} = 5.18$).[3] An examination of the suspicion probe indicated that no participants were aware of the experimental hypotheses.

Test of hypotheses. H_3 put forth competing predictions about the influence of the number of contact sources in producing consumer contamination. H_{3a} predicted that consumer product evaluations and purchase intentions would be lower if a product was thought to have had contact with many (versus one) consumers, whereas H_{3b} predicted that product evaluations and purchase intentions would be higher in this instance. A 2 (proximity to contact: medium versus far) × 2 (number of contact sources: one versus many) ANOVA with the evaluation index as the dependent variable produced a main effect for number of contact sources ($F(1, 77) = 4.45, p < .05$) and a main effect for proximity to contact ($F(1, 77) = 5.08, p < .05$). The interaction effect between the two independent variables was not significant ($F < 1$). Consistent with H_{3a}, participants evaluated the product more positively when they believed that only one other person had touched the product before them ($M_{one} = 3.86$) than when they believed that several other people had touched the product before them ($M_{many} = 3.19$). Furthermore, consistent with H_1, participants evaluated the product more positively when the product was located on a regular rack in the store ($M_{far} = 3.89$) than when it was on the return rack ($M_{medium} = 3.19$).

We conducted a similar analysis using WTP as the dependent variable. The results for WTP also produced two main effects (proximity to contact: $F(1, 77) = 4.38, p < .05$; number of contact sources: $F(1, 77) = 4.73, p < .05$). Participants were willing to pay more to purchase the product when it was located on the regular rack in the store ($M_{far} = \$20.00$) than when it was on the return rack ($M_{medium} = \$11.85$), and they were willing to pay more when only one person had touched it (M_{one} = \$16.18) than when multiple people had touched it (M_{many} = \$11.72). Again, there was no significant interaction effect ($p > .10$).

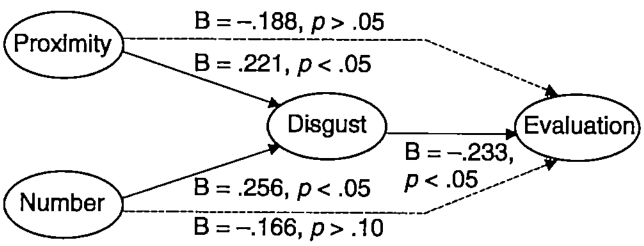
---

[3] We conducted a one-way ANOVA with the control group and the number of contact source conditions for the number index ($F(2, 100) = 15.38, p < .001$; $M_{control} = 2.15$; control versus one: $t(98) = 2.64, p = .01$; control versus many: $t(98) = 5.39, p < .001$).

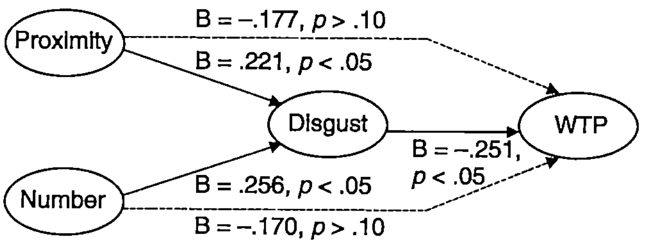
A two-way ANOVA with the disgust index as the dependent variable produced only two main effects (proximity to contact: $F(1, 77) = 4.24$, $p < .05$; $M_{medium} = 3.13$, $M_{far} = 2.55$; number of contact sources: $F(1, 77) = 5.66$, $p < .05$; $M_{one} = 2.52$, $M_{many} = 3.19$). To determine whether feelings of disgust led to changes in product evaluations, we conducted a mediation analysis. To test for mediation, we estimated a path analysis that included direct paths for the effects of proximity to contact and number of contact sources on product evaluations and indirect paths mediated by disgust. The results produced a significant indirect path, indicating that disgust fully mediated the effects of the independent variables on product evaluations (see Figure 4, Panel A). We also conducted a similar path analysis for WTP, and the results indicated that disgust fully mediated the impact of contagion cues on WTP (see Figure 4, Panel B). Finally, to test the causal relationship between evaluations and WTP, we conducted another path analysis. The results indicated that evaluations mediated the impact of contamination cues on WTP for the contaminated product (see Figure 4, Panel C).

FIGURE 4
Study 3: Path Analysis

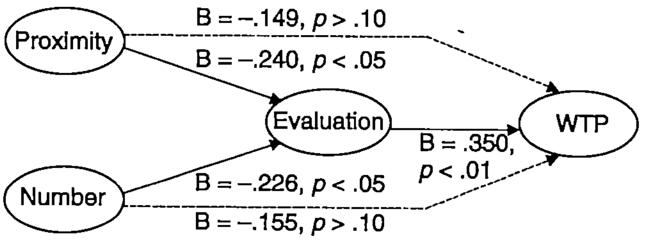
A: Mediating Role of Disgust on Evaluations



B: Mediating Role of Disgust on WTP



C: Mediating Role of Evaluations on WTP



Control group. We compared the control group with the number-of-contact-sources conditions in a series of one-way ANOVAs. The results were significant for all the key dependent variables (product evaluations: $F(2, 98) = 15.19$, $p < .001$; $M_{control} = 5.39$; WTP: $F(2, 98) = 14.90$, $p < .001$; $M_{control}$ = \$25.45; disgust: $F(2, 98) = 16.00$, $p < .001$; $M_{control} = 1.31$). Post hoc tests indicated that the control group differed significantly from the one-source and many-sources conditions for all the variables (product evaluations: control versus one [$t(98) = 3.92$, $p < .001$] and control versus many [$t(98) = 5.50$, $p < .001$]; WTP: control versus one [$t(98) = 3.76$, $p < .001$] and control versus many [$t(98) = 5.46$, $p < .001$]; disgust: control versus one [$t(98) = 3.72$, $p < .001$] and control versus many [$t(98) = 5.66$, $p < .001$]).

Discussion of Study 3

The findings of Study 3 provide additional evidence for our general proposition that perceived contact between a consumer and a product results in other consumers having less favorable evaluations of the touched product. Our results show support for H_{3a} and SIT, indicating that when consumers believe that many people have touched a product, their evaluations and WTP for the product are significantly lower. Rather than signaling the popularity of a product as H_{3b} predicted, multiple contact sources seemed to increase the level of contamination realized. Our findings are telling, given that the experimental procedure may have also signaled the popularity of the product (i.e., the item was "the only one left in stock"). The results for Study 3 also show consistent effects for proximity to contact such that participants responded more negatively to a product that was situated closer to a possible contact location. Finally, as in Studies 1 and 2, the results of Study 3 also show that the feelings of disgust that consumers experience in response to product contamination lead to the changes in product evaluations and WTP.

Conclusions and Implications

Recent marketing research has found that consumers like to use the sense of touch to learn about products (Mooy and Robben 2002; Peck and Childers 2003). However, the current research demonstrates that though consumers enjoy touching products themselves, they respond negatively when other consumers have had contact with the products. Using a retail shopping environment, the current research provides the first evidence for a theory of consumer contamination, in which consumers are believed to contaminate products simply by having contact with them. In general, we find that contamination cues that increase the salience of contact between another consumer and a product have a negative impact on consumer responses to the product.

Theoretical Implications

Contamination cues. Across three field experiments, we demonstrate that when proximity to contact is closer and when the number of contact sources is higher, consumer evaluations, purchase intentions, and WTP for a touched product decrease. Notably, the results for the impact of proximity to contact hold only when a short period of time has elapsed since the product was perceived as being touched. When time since contact is longer, the effects of consumer contamination appear to dissipate.

We argue that these contamination cues increase the salience of contamination by effectively signaling to the consumer that the product has previously come in contact with another person or people. It is important to recognize that in our research, this contact was implied and not physical in nature. As is shown by the null findings from our dirty index, although consumers were cognizant that no physical contamination was engendered to the product by the implied touching, they were still influenced by the contamination from the other consumer. Indeed, the lingering perceived essence of another seems to be enough to stigmatize the product for the consumer. The contamination cues we identify appear to activate a contamination heuristic, which, though arguably misapplied, has important repercussions for consumer evaluations and intentions.

Time and contamination. The identified temporary nature of consumer contamination emphasizes a key theoretical finding that builds on previous work in this area (Hejmadi, Rozin, and Siegal 2004; Rozin et al. 1989). Prior research has examined cases of contamination in which people had extremely strong attitudes toward the source of contamination. These sources have been life threatening (e.g., poison) or have been people about whom others have strong, preexisting negative attitudes and associations (e.g., Adolf Hitler). After the source has come into contact with the object, these strong negative attitudes toward the source seem to transfer permanently and are then associated with the object (Hejmadi, Rozin, and Siegal 2004; Rozin et al. 1989). In the current research, however, consumers have no preexisting attitudes toward the contamination source, because the source is either unfamiliar or completely unknown to consumers. Furthermore, consumers recognize that the contamination is not life threatening in nature. Thus, there is no permanent association to be transferred on contact, and any contamination effects that occur are indicative of a perceived transfer of essence. Our findings qualify previous thought by empirically showing that for the type of contamination investigated in the current research, an increase in temporal distance mitigates the strength of the contamination effect.

Feelings of disgust. Consistent with previous research (Rozin, Haidt, and McCauley 1993; Rozin, Markwith, and Nemeroff 1992; Rozin, Millman, and Nemeroff 1986), we found that disgust operates readily in the domain of contamination in the consumption context. In particular, the results indicate that consumers lower their evaluations of touched products because they feel disgusted by the contamination from other shoppers. The analyses demonstrating full mediation of disgust on evaluations, purchase intentions, and WTP support this assertion. Importantly, we show discriminate validity for disgust by finding null effects for other negative emotions. The visceral reaction that disgust represents in this context points to the basic nature and strength that contamination effects exemplify.

Managerial Implications

Store design and merchandising. The consumer contamination effects demonstrated in each of our three studies underscore the importance of management decisions related to store layout and design. Although it is critical to enable consumers to touch merchandise (Mooy and Robben 2002; Peck and Childers 2003), consideration should also be given to limiting how and where consumers touch products. Effective store design can provide opportunities and constraints with respect to consumer contamination. For example, instead of using product bins or table displays that encourage rummaging and searching, firms can effectively design shelving and rack displays to minimize consumer handling of products. Furthermore, a front-room/back-room layout that provides "fresh" products for trial can potentially mitigate consumers' contamination perceptions. Retail environments in which consumers try on merchandise (e.g., clothing, footwear) need to be sensitive to consumers viewing others handling the products. As *The Washington Post* (Pressler 2003) noted recently, effective design and management of dressing room areas is an understated factor in retailer success. Notably, the article indicates that the primary complaint of clothing shoppers is that other consumers tried on clothing that cluttered the dressing room.

Appropriate merchandising and well-managed product displays are also central in effectively addressing contamination concerns. Cues that a product has previously been used or touched, such as ripped tags or disheveled displays, raise contamination fears and negatively affect consumer evaluations. Managers must make an effort to minimize the salience of contamination so that consumers are not motivated to think that a given product has been contaminated. Effective merchandising of products can be useful in allaying and preventing contamination effects from materializing. Indeed, Underhill (2000, p. 179) notes that retailers "place the dark shades down where they can be reached easily, and display the beiges and pale grays and off-whites up high, where they can be seen but not touched. If a sweater displayed on a table comes in several shades, you'll always find the lighter ones on the bottom and the dark on top, where they'll be rubbed and grubbed, but who will know?"

Service personnel and communication. Our results also point to a strong need for proper employee training and service practices. Employees need to be cognizant of the issues surrounding consumer contamination and have the ability to address related consumer concerns. MAC cosmetics is an example of a company that has responded to this opportunity, turning contamination concerns into a competitive advantage in the industry. MAC employees are trained to wipe sample lipstick both before and after consumer trial and use disposable applicators (e.g., mascara, blush) for consumer promotion activities. Employees must be mindful to remove the evidence of consumer contamination as quickly as possible in a retail environment. Refolding, restocking, repackaging, and returning products quickly to shelf locations are examples of simple activities that can alleviate consumers'contamination concerns. Companies, such as Nordstrom and Banana Republic, that encourage their personnel to be vigilant in this regard communicate a safe, comfortable, and individualized shopping environment in which sales staff are sensitive to specific consumer needs and concerns about contamination.

In a broader sense, a company's communication with consumers should also be sensitive to contamination. Advertising, promotions, and other communication efforts should be careful not to raise contamination concerns in their creatives. Communication that highlights the incidence of consumer contact and identifies who is touching products may result in contamination signals that alienate potential customers.

Policy and procedures. Acceptance of the effects and implications of consumer contamination gives way to the effective integration of this issue into store systems and policies. Policies that are created specifically to respond to contamination outcomes provide direction and a means to manage the negative implications of this phenomenon. As we alluded to previously, effective policies and procedures that empower employees to recognize and address consumer contamination are an effective approach. The development of policies for more physical forms of contamination is also imperative for organizations. For example, Best Buy has adopted a classification of "Open Box Products," which describes products that have been handled, opened, or have had the packaging disrupted. In this case, the product and warranty are intact, but the policy calls for an "Open Box" labeling of the product and a price discount. Return policies, warranties, and consumer trial all should be examined with an eye toward contamination effects. Recent trends in the retail environment of bulimic shopping and retail borrowing (Piron and Young 2000) make this a growing concern.

Limitations and Future Research Opportunities

The limitations of our research are inherent to the experimental methodology and sample we used in conducting our studies. Each of our three studies used the same product (i.e., clothing) as the experimental stimuli, followed a similar procedure, and was consistent in the retail environment. However, as we mentioned previously, because contamination effects are expected to differ depending on the salience *and* intimacy of contact (Angyal 1941), the results reported herein may be limited only to product categories for which consumers have close, intimate contact. Indeed, it seems likely that contamination effects are moderated by the level of contact or intimacy that consumers have with a product while using it: The more contact a product has with a consumer's body, the greater are the contamination effects. Prior anthropological work supports this notion, demonstrating stronger contamination effects for ingestion, the most intimate form of contact, than for mere interaction with an object (Appadurai 1981). This suggests that contamination effects in retail contexts could be most prevalent for food items, creating a true challenge for grocery retailers. Further research should test contamination effects across different product categories to examine these issues more closely.

Another possible moderating factor of contamination effects is product packaging. In the studies we presented herein, products were always free from packaging and completely vulnerable to contact with consumers. Thus, it would be interesting to investigate whether product packaging can protect products from consumer contamination and whether the level of protection varies by the packaging material. Recent work that has examined contamination effects among products (Morales and Fitzsimons 2005) has found that clear packages facilitate contamination effects and opaque packages prevent them. Specifically, this research shows that when a consumable product is displayed inside a clear, plastic package and is touching a disgusting product, the consumable product becomes contaminated, but when the consumable product is displayed inside a solid, paper package, evaluations of the product remain constant. Whether these findings for packaging hold true for consumer contamination, however, remains unclear and is left for further research.

Another area for extension is to investigate the potential for positive contamination effects in a consumer context. As we mentioned previously, research has found that when a contamination source has a positive relationship with the recipient (i.e., friend, lover), responses toward contamination are more positive in nature than when a relationship does not exist. This result is consistent with young sports fans keeping a home-run ball that is hit by an admired baseball player or David Beckham's missed soccer ball selling for $16 million on eBay. However, unlike the cases we examined in the current work, these examples of positive contamination are primarily the result of a symbolic interaction model based on associations rather than a residue model resulting from physical contact alone (Nemeroff and Rozin 1994). As such, for positive contamination effects to occur in consumer settings, the identity of the contamination source needs to be known and have a strong, positive association with the consumer. In contrast to the negative contamination effects demonstrated here, simply knowing that a product has been touched by an anonymous source is not enough for positive contamination to arise. For positive effects, the contamination source should be as important as the contact.

In addition to influencing whether positive or negative contamination effects will occur, the source of contamination might also serve as a moderating factor in general. For example, salespeople are constantly arranging displays and moving around products. Do consumers view this touching as contamination? How do consumers integrate their rules for contamination across different "actors" (i.e., salespeople versus other shoppers) in the consumption context? Is touching by some consumers actually considered positive, whereas touching by other consumers considered extremely negative? Further research could attempt to understand how touching by anonymous versus different known sources fits into the developed theory of consumer contamination.

Similarly, additional work is also needed to examine how the type of retail outlet influences consumer contamination effects. Consumers might have different ideas for the amount of contamination they would expect to find in an upscale store such as Nordstrom versus a discount center such as Wal-Mart. It may be that some retail outlets generate strong contamination effects because physical contact among consumers and products is expected to be controlled and limited, whereas others do not because previous consumer contact with products is expected, as in a consignment store. Such speculation suggests that consumers may create rules (and exceptions) or frame information differently (Rozin and Nemeroff 1990) depending on the contamination context. These and other contextual differences could be investigated in further research on contamination.

Finally, in the current research, we consider the impact of three types of contamination cues: proximity to contact, time elapsed since contact, and number of contact sources. In the retail environment, there are likely several other types of contamination cues that may be prevalent. For example, physical clues (e.g., cut or handwritten price tags or marked products) related to contact may all have a role in the realization of contamination effects. Future investigation should broaden our research framework to provide additional insight into the antecedents and outcomes of consumer contamination.

REFERENCES

Alba, Joseph, John Lynch, Barton Weitz, Chris Janiszewski, Richard Lutz, Alan Sawyer, and Stacy Wood (1997), "Interactive Home Shopping: Consumer, Retailer, and Manufacturer Incentives to Participate in Electronic Marketplaces," *Journal of Marketing*, 61 (July), 38–54.

Angyal, A. (1941), "Disgust and Related Aversions," *Journal of Abnormal and Social Psychology*, 36, 393–412.

Appadurai, Arjun (1981), "Gastro-Politics in Hindu South Asia," *American Ethnologist*, 8 (3), 494–511.

Argo, Jennifer J., Darren W. Dahl, and Rajesh V. Manchanda (2005), "The Influence of a Mere Social Presence in a Retail Context," *Journal of Consumer Research*, 32 (September), 207–212.

Baron, Reuben M. and David A. Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (December), 1173–82.

Citrin, Alka Varma, Donald E. Stem Jr., Eric R. Spangenberg, and Michael J Clark (2003), "Consumer Need for Tactile Input: An Internet Retailing Challenge," *Journal of Business Research*, 56 (11), 915–22.

Frazer, J.G. ([1890] 1959), *The New Golden Bough: A Study in Magic and Religion*, abridged ed., T.H. Gaster, ed. New York: Macmillan.

Haidt, Jonathan, Clark McCauley, and Paul Rozin (1994), "Individual Differences in Sensitivity to Disgust: A Scale Sampling Seven Domains of Disgust Elicitors," *Personality and Individual Differences*, 16 (5), 701–713.

Hejmadi, Ahalya, Paul Rozin, and Michael Siegal (2004), "Once in Contact, Always in Contact: Contagious Essence and Conceptions of Purification in American and Hindu Indian Children," *Developmental Psychology*, 40 (4), 467–76.

Kotler, Philip and Murali K. Mantrala (1985), "Flawed Products: Consumer Responses and Marketer Strategies," *Journal of Consumer Marketing*, 2 (3), 27–36.

Latané, Bibb (1981), "The Psychology of Social Impact," *American Psychologist*, 36 (4), 343–56.

——— and Sharon Wolf (1981), "The Social Impact of Majorities and Minorities," *Psychological Review*, 88 (5), 438–53.

Liberman, Nira and Yaacov Trope (1998), "The Role of Feasibility and Desirability Considerations in Near and Distant Future Decisions: A Test of Temporal Construal Theory," *Journal of Personality and Social Psychology*, 75 (1), 5–18.

Mauss, Marcel ([1902] 1972), *A General Theory of Magic*, R. Brain, trans. New York: Norton.

McCabe, Deborah Brown and Stephen M. Nowlis (2003), "Effect of Examining Actual Products or Product Descriptions on Consumer Preference," *Journal of Consumer Psychology*, 13 (4), 431–49.

Meigs, Anna S. (1984), *Food, Sex, and Pollution: A New Guinea Religion*. New Brunswick, NJ: Rutgers University Press.

Mooy, Sylvia C. and Henry S.J. Robben (2002), "Managing Consumers' Product Evaluations Through Direct Product Experience," *Journal of Product and Brand Management*, 11 (6–7), 432–45

Morales, Andrea C. and Gavan J. Fitzsimons (2005), "Product Contagion: Changing Consumer Evaluations Through Physical Contact with 'Disgusting' Products," unpublished manuscript, Marketing Department, Arizona State University.

Nemeroff, Carol J. (1995), "Magical Thinking About Illness Virulence: Conceptions of Germs From 'Safe' Versus 'Dangerous' Others," *Health Psychology*, 14 (2), 147–51.

——— and Paul Rozin (1994), "The Contagion Concept in Adult Thinking in the United States: Transmission of Germs and of Interpersonal Influence," *Ethos: Journal of the Society for Psychological Anthropology*, 22 (2), 158–86.

O'Reilly, Lynn, Margaret Rucker, Rhonda Hughes, Marge Gorang, and Susan Hand (1984), "The Relationship of Psychological and Situational Variables to Usage of a Second-Order Marketing System," *Journal of the Academy of Marketing Science*, 12 (3), 53–76.

Peck, Joann and Terry L. Childers (2003), "To Have and To Hold: The Influence of Haptic Information on Product Judgments," *Journal of Marketing*, 67 (April), 35–48.

Piron, Francis and Murray Young (2000), "Retail Borrowing: Insights and Implications of Returning Used Merchandise," *International Journal of Retail & Distribution Management*, 28 (1), 27–36.

Pressler, Margaret W. (2003), "Trying Experiences: Tiny and Untidy Fitting Rooms Can Give Shoppers Fits," *The Washington Post*, (November 23), F05.

Rozin, Paul and April E. Fallon (1987), "A Perspective on Disgust," *Psychological Review*, 94 (1), 23–41.

———, Jonathan Haidt, and Clark McCauley (1993), "Disgust," in *Handbook of Emotions*, Michael Lewis and Jeannette M. Haviland, eds. New York: Guilford Press, 575–94.

———, Maureen Markwith, and Clark McCauley (1994), "Sensitivity to Indirect Contacts with Other Persons: AIDS Aversion as a Composite of Aversion to Strangers, Infection, Moral Taint, and Misfortune," *Journal of Abnormal Psychology*, 103 (3), 495–504.

———, ———, and Carol Nemeroff (1992), "Magical Contagion Beliefs and Fear of AIDS," *Journal of Applied Social Psychology*, 22 (14), 1081–1092.

———, Linda Millman, and Carol Nemeroff (1986), "Operation of the Laws of Sympathetic Magic in Disgust and Other Domains," *Journal of Personality and Social Psychology*, 40 (4), 703-712.

——— and Carol Nemeroff (1990), "The Laws of Sympathetic Magic: A Psychological Analysis of Similarity and Contagion," in *Cultural Psychology: Essays on Comparative Human Development*, James E. Stigler, Richard A. Shweder, and Gilbert Herdt, eds. New York: Cambridge University Press, 205–232.

———, ———, Marica Wane, and Amy Sherrod (1989), "Operation of the Sympathetic Magical Law of Contagion in Interpersonal Attitudes Among Americans," *Bulletin of the Psychonomic Society*, 27 (4), 367–70.

Priya Raghubir & Eric A. Greenleaf

Ratios in Proportion: What Should the Shape of the Package Be?

Consumers' reactions to rectangles have implications for package and product design. In two lab studies and an analysis of field data, the authors find that the ratio of the sides of a rectangular product or package can influence purchase intentions and preferences and is related to marketplace demand. In more exploratory inquiries, the authors also find that the impact of this ratio on consumers depends on the relative seriousness of the context in which the product is used. Furthermore, ratio can also affect product perceptions, and consumers appear to prefer a range of contiguous ratios for different contexts rather than a particular ratio.

One of the oldest controversies in aesthetics, extending back to the ancient Greeks, is whether people prefer certain ratios for the sides of rectangles. Investigations into this question began in the 1870s with Fechner's work and continue to the present day. The results of this research have been mixed. The research has not produced definitive conclusions about whether people consistently prefer particular rectangular ratios (Benjafield 1985; Green 1995).

Rectangular preferences have important implications for marketers because so many products and packages are rectangles. Recent research in the trade press suggests that a package's shape is a critical way for a brand to differentiate itself because package design can affect consumers' purchase decisions (Sherwood 1999). This has led many firms to focus on product and package shape, and small changes in package shapes can have a large influence on sales and profits (Prince 1994). Many of the "new age" beverages of the 1990s (e.g., Snapple, Arizona Iced Tea, Perrier) paid close attention to bottle design and are credited with helping spark a packaging renaissance (Miller 1994). With so much attention being given to the shape of products and packages, brand managers and designers who are responsible for introducing new products or for creating new shapes and packages for existing products need guidance on how rectangular proportions can affect consumer purchase intentions, preferences, and demand.

Marketers have proposed conceptual models of factors that may affect consumer reaction to product form (Bloch 1995). Bloch's (1995) model proposes antecedents, mediators, and moderators of the impact of aspects of package shape on consumer preferences, intentions, and choice. Although Bloch's model is rich in insights into possible influences on consumers' reactions to product form and how consumers process visual information, it is primarily a conceptual model. Research has also begun to examine how package shape affects the managerially important variables of volume perceptions and consumption (e.g., Wansink and Van Ittersum 2003).

The current article examines one particular aspect of product design: We take the concept that people have different preferences for different rectangular ratios, an area that has received considerable attention in psychology and aesthetics, and apply it empirically to the consumer realm to answer questions of interest to marketing managers and product designers who deal with rectangular products and packages. Our principal objective is to investigate whether the ratio of the sides of a rectangular product or package—in other words, rectangles of interest to marketers—can affect consumer purchase intentions and preferences and to examine how ratio is related to marketplace demand. An additional objective is to examine, on an exploratory basis, additional issues in this domain of interest to marketing managers, such as whether ratios affect product perceptions, whether the impact of ratio on purchase intentions and preferences and the ratio's relationship to market share depend on the context in which the rectangular product is used, and whether consumers tend to prefer a contiguous range of rectangular ratios for products and packages rather than one particular ratio.

Prior Research on Preferences for Rectangles

Aestheticians have long been interested in whether people find rectangles with certain ratios to be more pleasing than other ratios and how a rectangle's aesthetic qualities might be related to the mathematical properties of its ratio (Alberti [1485] 1988; Huntley 1970; Lawlor 1982). This line of inquiry has also examined how often particular ratios appear in nature, such as in the structure of plants and animals (Ghyka 1977; Lawlor 1982).

Priya Raghubir is Associate Professor of Marketing, Haas School of Business, University of California at Berkeley (e-mail: raghubir@haas.berkeley.edu). Eric Greenleaf is Associate Professor of Marketing, Stern School of Business, New York University (e-mail: egreenle@stern.nyu.edu). The authors contributed equally to this article. The authors thank Nikki Lee for her assistance with stimuli preparation and data entry, Suchi Chandran for her research assistance, and Rishi Chand and Jane Gu for their assistance with survey data collection. The authors gratefully acknowledge the comments of Bob Krider on a previous draft of this manuscript and the comments of participants at the IC-1 conference on Visual Marketing in Ann Arbor, Michigan, where a previous version of this article was presented.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 70 (April 2006), 95–107

Much of this work has focused on the "golden ratio," or the Φ (phi) rectangle, which has a ratio of approximately 1.618. The number Φ has numerous special mathematical properties that have been studied by mathematicians such as Pythagoras, Euclid, and Fibonacci (Herz-Fischler 1987; Livio 2002). Some of these properties translate into visual properties (see Ghyka 1977; Huntley 1970; Livio 2002). The Φ is the only rectangle for which the ratio of the longer side to the sum of the sides equals the ratio of the short side to the longer side. If a square is removed from a Φ rectangle, the remaining rectangle is still a Φ rectangle. Repeating this process several times yields an attractive spiral such as that found in nautilus shells.

Advocates of the golden ratio's aesthetic appeal point to its alleged use in major works of art and architecture, such as the building blocks of the Great Pyramids, the Parthenon's facade, and paintings by Georges Seurat (see Ghyka 1977; Huntley 1970), whereas others are cautious about such claims (Livio 2002). In Bloch's (1995) theoretical review of the factors that affect "ideal product form," he posits that the golden ratio may appeal to people because of overall "gestalt" qualities that supersede any one particular characteristic. Although Bloch acknowledges the differences just discussed, he cites the gestalt qualities in "enduringly popular architectural forms" (p. 22) and notes that "it could prove useful for managers involved with design decisions to be acquainted with Gestalt principles ... [and that] firms may find it useful to include these considerations" (p. 22).

Scientific investigations of whether people actually prefer certain rectangles began with Fechner's (1871, 1876; for a recent translation, see Fechner 1997) pioneering work, which found a preference for the Φ rectangle. Other early researchers replicated Fechner's results (Lalo 1908; Thorndike 1917; for a review, see Green 1995). However, research supporting a preference for the Φ rectangle, including Fechner's, has been critiqued on methodological grounds (for reviews, see Green 1995; Plug 1980). Preferences can be influenced by participant instructions (Hekkert, Lieke, and Van Wieringen 1994), preference measures (McManus 1980; Ohta 1999), elicitation (Höge 1995, 1997), the range of ratios used and the location of a ratio in that range (Benjafield and McFarlane 1997; Godkewitsch 1974; Piehl 1976), whether rectangles have equal area (Benjafield 1976), use of average versus individual preferences (McManus and Weatherby 1997; Plug 1980), and whether the rectangle is part of a more complex pattern (Boselie 1984).

Our studies control for many of these methodological concerns. Furthermore, rather than clearly linking the studies' purposes to rectangular ratios, we use an indirect approach that is less likely to create demand artifacts (Green 1995; Höge 1995, 1997; Plug 1980). Study 1 uses a between-subjects rather than a within-subjects design; to disguise the study's purpose, Study 2 participants make additional purchase intention ratings and rankings for design alternatives that do not involve rectangles; and Study 3 uses actual marketplace sales.

Rectangles as Objects Used in a Particular Consumer Context

Most studies of rectangular preferences have examined rectangles as abstract shapes, devoid of any particular meaning or use. This limits their application to real-world stimuli (Koneni 1997). A major limitation for marketing is that consumers are not concerned with buying abstract rectangles but rather with products and packages that have a rectangular shape. Marketing managers and product designers need to understand how consumers are influenced by the proportions of these rectangular products and packages and use this knowledge to guide their own decisions about packaging and product dimensions. Bloch's (1995) conceptual model recognizes the importance of this issue and approaches the impact of product form in a consumer setting rather than treating form abstractly.

Thus, there is a need for empirical research that examines consumer reactions to rectangles as they are applied to products and packages, not just rectangles in the abstract sense. Because marketers' concerns extend beyond preferences to include consumers' purchase intentions and demand, constructs that have little meaning when applied to abstract rectangles, we examine these latter issues. Thus, we investigate consumer reactions to rectangular products and packages, not abstract rectangles without context. Previous work in marketing has examined the distribution of the ratios of rectangular print advertisements and how the placement of headlines, pictures, text, and brand logo within an advertisement affects preferences for the advertisement (DeHeer and Nobbe 2000). However, we do not know of previous work that has examined how the rectangular ratios themselves affect purchase intentions, product preferences, and market share for products.

The remainder of this article proceeds as follows: Study 1 tests the hypothesis that in a consumer setting, the ratio of a rectangle can affect consumers' purchase intentions and preferences for a particular product. On an exploratory basis, this study also tests whether ratio also affects certain consumer product perceptions. Study 2 replicates the test of this hypothesis for a different product category. On an exploratory basis, it also examines whether these purchase intentions depend on whether the context in which a product is used is relatively serious or relatively frivolous and whether higher purchase intent exists for a contiguous range of ratios rather than a particular ratio. Study 3 tests whether consumer demand for frequently purchased goods sold in rectangular boxes is related to rectangular ratio and whether the context in which a product is used moderates this relationship.

Study 1: Impact of Rectangular Ratio on Purchase Intentions, Preferences, and Perceptions

The Impact of Rectangular Ratio on Purchase Intentions and Preferences

Although the evidence that people prefer particular rectangles has been mixed, consistent throughout this research is the notion that preferences indeed differ across ratios; that is, participants in these studies do not favor all rectangles equally. Applied to the realm of consumer products, this suggests that rectangular ratio can affect consumer preferences and purchase intentions because both constructs are related to demand.

> H_1: The ratio of a rectangle's sides affects consumers' (a) purchase intentions and (b) preferences for rectangular products.

Design and Procedure

Our sample consists of 108 undergraduate students (56 females and 52 males) enrolled in an introductory marketing course; they completed the experiment for partial course credit. Participants were told that the experiment was a study of musical tastes among students at their university. They were asked to assume that they had been mailed invitations with descriptions of two concerts (Concert A and Concert B). The words in the description of Concert A, titled "Music of Indonesia: Songs Before Dawn," were printed in a square box that was six inches on each side (a 1:1 ratio). We used this concert as a control to create a reference point that people would compare with their judgments of a target concert (Concert B) and to control for individual differences in scale response. The target description of Concert B, titled "Dream Songs: Sounds of the Malaysian Rainforest," was printed in a horizontal (width > height, as opposed to a vertical rectangle [width < height]) rectangular box with the ratio of sides either 1.38 or 1.62 (the golden ratio). Both boxes had the same area and contained the same description of Concert B (see Figure 1). We decided to include the golden ratio because this rectangle has been the subject of so much inquiry.

Purchase intentions for CDs. To measure purchase intentions for a product closely related to the concert, we asked participants to rate their likelihood of buying a CD of the music for Concerts A and B on a seven-point scale (1 = "definitely not," 7 = "definitely").

Relative preference for concerts. To measure relative preference for attending Concert A versus Concert B, we asked participants to indicate which concert they would prefer to attend if the two concerts were held at the same time, on a six-point scale with intervals labeled "strongly," "moderately," and "slightly" prefer Concert A or B. Higher numbers reflect a preference for Concert B.

Covariates and controls. Because purchase intention for the CD of Concert B is one of the dependent variables of interest and because ownership of CDs is also a useful indicator of experience with music, participants answered questions about the number of CDs they owned for use as a potential covariate; the response categories were as follows: 10 or fewer = 1, 11–50 = 2, 51–100 = 3, 101–200 = 4, and more than 200 = 5. To measure tendency to attend concerts, which might also be a useful covariate in concert preferences, participants reported how many live music concerts they attended in an average year. Participants also reported their gender (female = 0, male = 1). To assess whether ratio affected a cognitive phenomenon—namely, depth of processing—after a filler task, we administered a surprise recall task that was followed by a five-item multiple-response "quiz" on the content of the information provided for Concert B.

Results

Purchase intentions. To examine whether the ratio of the target concert description affected purchase intentions for the CD of Concert B, we used an analysis of covariance (ANCOVA) with purchase intentions as the dependent variable. The independent variables were the two different ratios of the Concert B invitation and gender, both coded as a categorical variable, and the scale for the number of CDs owned. The adjusted mean purchase intention for the CD of Concert B when the description was in the 1.62 ratio (M = 3.05) was greater than that for the 1.38 ratio (M = 2.22; F(1, 103) = 6.22, $p < .05$). Gender and CD ownership were not significant (Fs < 1.0). Thus, the results support the impact of ratio on purchase intentions (H_{1a}). In this example, the 1.62 ratio created higher purchase intentions for the CD of Concert B than did the 1.38 ratio.

We also examined whether the impact of ratio on purchase intentions depended on an interaction between ratio and gender by adding this interaction term to the ANCOVA. This interaction term was not significant (F < .01).

Impact of rectangular ratio on consumer perceptions. Research on rectangular preferences in aesthetics and psychology has not usually examined attribute perceptions, because this work has tended to focus on abstract rectangles, not on consumer products that can be broken down into attributes. Because product perceptions matter to marketers, we believe that it is worth investigating, on an exploratory basis, whether the proportions of a rectangle in the context of a product or package affect consumers' perceptions of that product.

Although the aesthetic literature on rectangular preferences has not formally tested hypotheses that connect particular perceptions, this literature often discusses the importance of perception in the creation of these preferences. Commentators mention that preferred rectangles promote perceptions of harmony, proportion, balance, rationality, and mathematical beauty (Ghyka 1977; Huntley 1970; Lawlor 1982; Livio 2002; Svensson 1977 [who uses a line partitioning task]). These perceptions are also relevant to many, though by no means all, consumer products. To investigate the potential impact of ratio on these perceptions, we also included in Study 1 perceptual measures of the concerts that we believed were related to these aesthetic perceptions in the literature. We measured relative perceptions of Concert B and Concert A on scales that included the items "harmonious," "stable," "beautiful," "exciting," and "fast" (1 = "B much less than A," 4 = "B about the same as A," and 7 = "B much more than A"). We intended the last two to be reversed. We call this construct "harmony" to indicate its source in the literature on rectangular preferences, which often mentions harmony. However, we do not claim that these perceptions encompass all perceptions that may be relevant to rectangular products or packages or that harmony is always the dominant perception.

FIGURE 1
Stimuli for Study 1

Control Concert A: Ratio = 1:1

Music of Indonesia: Songs Before Dawn
"Compelling ... Fascinating musical smorgasbord from this cultural crossroads."

In the Banyuwangi region, located in the eastern end of Java, a vibrant and earthy musical genre called gandrung is performed. This tradition begins sometime around 9 p.m. and ends just before dawn. An unmarried female singer performs a beautiful suite of songs backed by a small ensemble of musicians who play violins and drums. The male guests at this dance pay money for the privilege of dancing with this mesmerizing songstress. Reticent yet dynamic, this call-and-response singing is accompanied by compelling rhythms played on percussion instruments that express and control the performance of the singers and dancers. This performance is by one of the music's finest living singers, Gadrung Temu. A recent review in the *SF Chronicle* called this "An enjoyable example of thoroughly contemporary and unusual exotica."

Target Concert B: Ratio = 1:1.38

Dream Songs: Sounds of the Malaysian Rainforest
"Rhythms that lure you irresistibly away to the land of dreams."

The Temiar people of the central Malaysian rainforest are deeply spiritual. Inspired by their vibrant natural environment, they have developed a musical landscape that evolves from their dreamsongs. Dreamsongs form the basis for community-wide singing and trance-dancing ceremonies, which celebrate marriages in Temiar society Sophisticated and delicate, this music contains cherished and exquisite compositions performed with an enchanting mix of xylophones, gongs, cymbals, fiddles, guitars, and breathtaking vocals. Instrumental and vocal music of the Temiar draws listeners into a realm of ornate tonal variations and textured rhythms. The performance features Kam Maitrasangi, a master of the Temiar dreamsong technique. A recent review in the *SF Chronicle* called this "A good introduction to this highly passionate and expressive artform."

Target Concert B: Ratio = 1:1.62

Dream Songs: Sounds of the Malaysian Rainforest
"Rhythms that lure you irresistibly away to the land of dreams."

The Temiar people of the central Malaysian rainforest are deeply spiritual. Inspired by their vibrant natural environment, they have developed a musical landscape that evolves from their dreamsongs. Dreamsongs form the basis for community-wide singing and trance-dancing ceremonies, which celebrate marriages in Temiar society. Sophisticated and delicate, this music contains cherished and exquisite compositions performed with an enchanting mix of xylophones, gongs, cymbals, fiddles, guitars, and breathtaking vocals. Instrumental and vocal music of the Temiar draws listeners into a realm of ornate tonal variations and textured rhythms. The performance features Kam Maitrasangi, a master of the Temiar dreamsong technique. A recent review in the *SF Chronicle* called this "A good introduction to this highly passionate and expressive artform."

Results for perceptual ratings. The five perceptual ratings all loaded on a single principal component (largest eigenvalue = 2.58, second eigenvalue = .84); absolute loadings ranged from .66 to .83. As we expected, "exciting" and "fast" loaded in the opposite direction of the other three perceptual scales. After reversing these two items, we averaged the five items to create a harmony index (α = .75); higher values indicate perceptions of greater harmony for Concert B than for A.

Relationship between perceptions and ratio. To examine the impact of ratio on the perceptual measure, we performed an ANCOVA with the perceptual scale as the dependent variable. The independent variables were the two categorical variables (ratio and gender) and CD ownership. The adjusted means show that participants rated Concert B as higher on the perception measure when the concert description used the ratio 1.62 (M = 4.70) than when the concert description used the ratio 1.38 (M = 4.32; F(1, 100) = 3.89, p = .051). The coefficient for gender was not significant (F < 1), nor was CD ownership (F = 2.57, p = .11). Thus, we find initial support that rectangular ratio can influence product perceptions, at least for a set of perceptions that are related to those often mentioned in the literature on rectangular preferences. We also examined whether gender and ratio interacted to influence perceptions by adding this interaction to the ANCOVA. However, the interaction term was not significant (F = .01, $p > .5$).[1]

Impact of ratio on relative preferences. We also examined the impact of ratio on relative preferences to attend Concert A versus Concert B; we included the potential role of perceptions as a mediator in this relationship. We conducted an ANCOVA on the relative preference of the concerts using the likelihood of attending the concert as the covariate and ratio and gender as the two categorical independent variables. We included attendance to control for individual differences in the propensity to attend live concerts. The analysis also included the perception measure. As we discussed previously, this step determines whether perceptions are a potential mediator of the impact of ratio on relative concert preferences.

Preference for Concert B versus Concert A was marginally greater when the description of the former was presented in the ratio of 1.62 (adjusted M = 4.15) than when it was presented in the ratio of 1.38 (adjusted M = 3.68; F(1, 99) = 2.88, p = .093). Thus, the direction of preference is the same as that for purchase intentions, in support of H_{1b}, though the evidence is not as strong as it is for CD purchase intentions (H_{1a}). Gender, concert attendance, and the perceptual measure did not have a significant relationship to preference, suggesting that perceptions did not mediate the relationship between ratio and preferences. The interaction between gender and ratio was not significant (F(1, 99) = .23, p = .63).

Impact of ratio on cognitions. There were no differences in the amount of recall of the concert information across conditions or in the number of questions that participants correctly answered in the surprise quiz task at the end of the procedure. Thus, we found no evidence for differential depth of processing in the two Concert B ratios.

Discussion

In summary, the results of Study 1 show that a difference in the ratio of the sides of a rectangular concert description can affect purchase intentions for the CD of music from the concert and, to a weaker extent, preferences. On an exploratory basis, we also found initial support that the ratio can affect product perceptions that are related to an index of perceptions discussed in the aesthetic literature on rectangles.

Study 2: Purchase Intentions for Invitation Cards

Study 2 has three purposes: First, we test H_1 in additional domains. Second, we examine H_{1a}, regarding purchase intentions, across a larger set of rectangular ratios than the two ratios used in Study 1. Third, on an exploratory basis, we examine whether the impact of rectangular ratios on purchase intentions depends on the relative seriousness versus frivolity of the consumer context in which the product is used. We also examine whether consumers have higher purchase intentions when rectangles are within a contiguous range of ratios or when this is only a single ratio.

Test of H_1 Across a Broader Set of Products and Rectangles

Although the comparison of two ratios in Study 1 creates a good starting point, many studies of rectangular preferences have included a larger set of ratios. Because rectangular consumer products and packages are found in the marketplace in a large range of ratios, we believe that it is useful to test H_{1a} with more rectangles from a larger ratio range. Study 2 examines nine ratios, ranging from the square to the 2:1 rectangle. Furthermore, in Study 2, consumers report purchase intentions for four different product variants of a rectangular product, along with their favorite and least favorite ratio. This method allows for within-person and across-person analysis of the impact of rectangular ratio, an approach recommended in prior research (Plug 1980).

Context Effects in Consumer Reactions to Rectangles

On an exploratory basis, Study 2 also examines whether consumer reactions to rectangular products and packages depend on the context in which the product is used.[2] Previous conceptual work on consumer reactions to product form (Bloch 1995) has proposed how context, such as situational factors and cultural and social context, may affect consumer reactions to product form. Thus, we propose a particular context that should be relevant for rectangular products and packages, namely, whether the product is used in a relatively more serious versus less serious context. We focus on this context effect for the following reasons: As we discussed previously, many aesthetic theories on preferences for particular rectangles, as well as other shapes, have emphasized that certain rectangles have properties such as harmony, proportion, balance, rationality, and mathematical beauty (Ghyka 1977; Huntley 1970; Lawlor 1982; Livio 2002). When applied to a consumer perspective, this rationale suggests that consumers should prefer these rectangles when the product is used in a context that makes the consumer more sensitive to and appreciative of these particular properties. When consumers are in a relatively serious context, they are more likely to appreciate these qualities than when they are in a relatively less serious, frivolous context. Whereas many consumer contexts are serious, such as going to an art museum, others are more frivolous and may even be intended to reduce harmony, balance, rationality, and so forth.

---

[1] We examined whether perceptions mediated the relationship between ratio and purchase intentions for the Concert B CD. A condition for mediation is that the proposed mediating variable must have a significant relationship to the dependent variable when the initial independent variable is included in the analysis (Baron and Kenny 1986). However, including the perceptions in the previous ANCOVA did not yield a significant effect for perceptions (F(1, 99) = .13, p = .72).

[2] Our use of the term "context" differs from that in the work of Benjafield and McFarlane (1997), who use it to describe how preferences for the Φ rectangle are affected by its location within the range of all rectangles that participants view.

We propose that in these latter, less serious contexts, products and packages that use ratios with these mathematical and aesthetic properties will not create higher purchase intentions than products and packages that use different ratios. We still expect purchase intentions to depend on ratio in relatively frivolous contexts, but we expect that this relationship differs from that in relatively serious contexts. Because we do not hypothesize which ratios will create the highest purchase intentions in relatively less serious contexts, we consider this investigation exploratory in nature. To pursue this line of inquiry, we include both relatively more serious and relatively less serious contexts in the product variants in Study 2.

Design and Procedure

The product of interest in Study 2 was custom-printed invitations for different events. The instructions read as follows: "Imagine that you are selecting invitations at the printer's shop. The printer tells you that you have to make three different decisions to design the invitations for each occasion." A total of 122 undergraduate students who were enrolled in a marketing course completed the experiment for partial course credit.

We manipulated rectangular ratio by asking participants to provide purchase intentions for rectangular invitation blanks for each of four occasions. The nine blanks varied in ratio from 1:1 to 2:1 in steps of approximately .13, which included the 1.62 (Φ) rectangle. Although Study 2 was not intended to focus specifically on the Φ ratio, we avoided placing the Φ ratio in the middle of the range because this placement can affect preferences (Benjafield and McFarlane 1997).

We labeled the card blanks with the letters A–I in the bottom right-hand corner using four different random orders; these labels were written to give half of the participants horizontal invitations and half of the participants vertical invitations. Although eye-tracking studies of how people view quadrangles have shown that scanning patterns do not vary consistently with quadrangle type, including differences in horizontal versus vertical orientation (Ohta 1999), we believe that it is advisable to balance this orientation. Card blanks were shuffled into random order before being placed in an envelope. Thus, there were 8 different types of envelopes used (4 order of letters × 2 orientations) × 4 (order of occasions). To increase task realism and disguise the study's purpose, participants made two other design decisions. The decisions were to choose envelope color (white, blue, green, or red) and the invitation's font (there were five different options).

The invitation contexts. To examine whether ratio interacts with context to influence purchase intent, each participant was asked to choose cards for four different contexts, two of which were relatively more serious and two of which were relatively less serious. In a pretest ($n = 40$), we evaluated relative seriousness using a seven-point scale, anchored by "not at all serious" and "very serious." The occasions chosen to be relatively more serious, which were a piano recital of classical piano music ($M = 4.91$) and a business presentation for prospective customers for accounting software ($M = 5.24$), received higher seriousness ratings than the occasions chosen to be relatively less serious, which were a three-year-old child's birthday party ($M = 1.92$) and an outing with adult friends to a three-ring circus ($M = 1.68$); participants rated both relatively serious contexts as more serious than the relatively less serious contexts ($p < .05$).

Measures of purchase intent. We measured purchase intent for each invitation blank for each of the four occasions in two ways: First, participants indicated the likelihood that they would choose each card option (A–I) on a six-point scale, anchored by "very likely" (6) and "very unlikely" (1). Because critiques of rectangular preference research find that rankings and ratings are complementary, participants also indicated which of the nine card blanks they were "most likely to choose" and "least likely to choose." Rankings preserve interitem distances, whereas ratings are less sensitive to aggregation and scale usage effects (Green 1995; Plug 1980).

Measure of relative seriousness of the four occasions. Because the relative seriousness of consumer occasions has not been measured as a construct, we could not use an established scale for this purpose. Accordingly, we generated five items to measure this construct. After providing purchase intentions, participants rated each of the four occasions using five seven-point semantic differential scales, anchored by "frivolous/serious," "orderly/disorderly," "unified/fragmented," "volatile/stable," and "rational/emotional." We reverse-coded the first and fourth variables (frivolity and volatility) so that final item scores ranged from 1 ("relatively more serious") to 7 ("relatively less serious").

Results

Manipulation check of context manipulation. A factor analysis for each of the four occasions confirmed that the five variables in the measure of relative seriousness loaded onto a single factor for each occasion (eigenvalues = 2.28, 2.86, 2.23, and 2.34 for birthday, accounting software, circus, and piano concert, respectively). We averaged the five variables to create an index of relative seriousness for each of the four occasions. This measure displayed a reasonable level of reliability for a new scale (α = .69, .79, .68, and .66 for birthday, accounting software, circus, and piano concert, respectively). A repeated measures 2 × 2 (context × replicate) analysis of variance (ANOVA) on this measure showed that participants rated the two occasions intended to be relatively more serious as more serious (accounting = 1.65; piano concert = 2.56) than the occasions intended to be relatively less serious (circus = 4.92; birthday party = 5.07; main effect of context: F(1, 121) = 895.25, $p < .001$). The two occasions within each context domain differed from each other (main effect of replicate: F(1, 121) = 28.80, $p < .001$), and this difference was greater for the two relatively more serious occasions (interaction: F(1, 121) = 70.04, $p < .001$).

Results for purchase intentions. To test H_{1a}, we first analyzed the purchase intention ratings using a 9 × 2 × 2 (ratio × context × replicate) within-subjects, repeated measure ANOVA. As H_{1a} predicted, intent ratings for cards differed by ratio (main effect: F(8, 848) = 64.18, $p < .001$). Mean purchase intention ratings for each card and ratio appear in Figure 2.

The rankings data also support H_{1a} (see Table 1). For each occasion, the number of participants who ranked each card as their first choice was contingent on the ratio of the cards ($\chi^2(8)$ = 73.56, 48.45, 30.15, and 68.63 for accounting, piano, circus, and birthday, respectively; $ps < .01$). The number of participants rating each card as their least favorite was also related to ratio ($\chi^2(8)$ = 164.29, 118.13, 105.07, and 180.44 for accounting, piano, circus, and birth-

FIGURE 2
Study 1: Mean Purchase Intention Ratings for Cards, by Ratio and Context

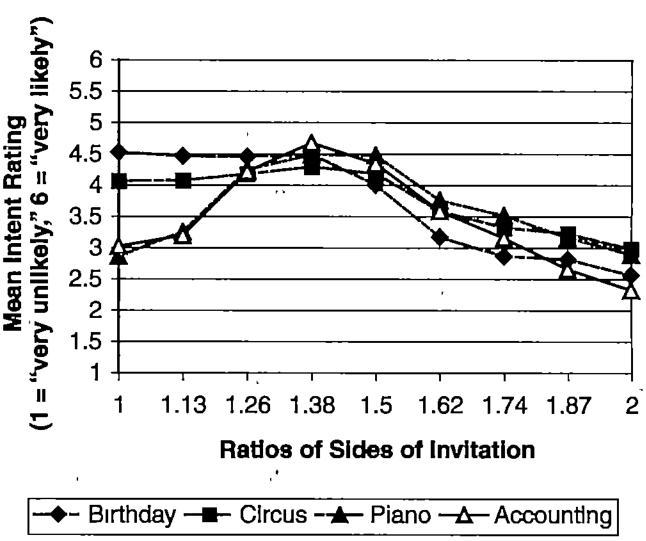


TABLE 1
Study 1: Purchase Intentions for Rectangular Cards by Ratio of Sides (Number of Participants Ranking a Card as First or Last Choice)

	Relatively More Serious Contexts			Relatively Less Serious Contexts		
	Accounting	**Piano**	**Total**	**Circus**	**Birthday**	**Total**
Number of First-Ranked Choices						
1:1	10	7	17	23	34	**57**
1.13	8	8	16	17	18	**35**
1.26	26	18	44	19	25	44
1.38	30	32	62	22	14	**36**
1.5	29	24	53	13	15	**28**
1.62	10	13	23	7	4	**11**
1.74	3	9	12	2	2	**4**
1.87	4	5	9	8	3	11
2	4	8	12	12	8	20
Number of Last-Ranked Choices						
1:1	31	28	59	13	7	**20**
1.13	8	17	25	13	5	18
1.26	4	2	6	9	8	**17**
1.38	0	3	3	3	3	6
1.5	3	4	7	6	5	11
1.62	7	3	10	6	7	13
1.74	4	8	12	12	11	**23**
1.87	16	15	31	13	18	31
2	51	44	95	48	59	107

Notes: Numbers in bold indicate a frequency that is significantly different across the two seriousness contexts at $p < .05$.

day, respectively; *ps* < .01). Thus, the purchase intention ratings and the first-choice and last-choice rankings all support H_{1a}.

Impact of context on purchase intentions. Next, we examined whether purchase intentions for the different rectangular cards depended on whether the card was being chosen for a relatively serious or frivolous occasion. In an ANOVA of the purchase intention ratings, the main effect of ratio was contingent on context (ratio × context interaction: $F(8, 848) = 13.91$, $p < .001$). The overall analysis also produced additional significant effects of context ($F(1, 106) = 15.73$, $p < .001$), a replicate × context interaction ($F(1, 106) = 4.01$, $p < .05$), and a significant three-way interaction ($F(8, 848) = 8.02$, $p < .001$). The interaction effects involving the replicate factor reflect that the two relatively more serious contexts differed more in their degree of rated seriousness than did the two relatively less serious contexts. The latter effect is not of concern to our predicted effects, and thus we do not analyze it further.

The number of participants who ranked each ratio as their first choice was also contingent on the four occasions (overall $\chi^2(24) = 71.83$, $p < .001$), as was the proportion of participants who ranked each ratio as their least favorite choice (overall $\chi^2(24) = 47.60$, $p < .001$). Thus, across the three measures of purchase intentions, first choice, and worst choice, the results suggest that preferences for the different ratios of the cards depended on the relative seriousness of the occasion for which the invitation was being selected.

Preference for a range of ratios rather than a particular rectangle. Prior research has often found that people prefer a contiguous range of rectangles rather than a particular ratio (Benjafield 1976; McManus 1980; Piehl 1976, 1978; Svensson 1977). If this finding applies in a consumer context, it could be important to marketers and brand managers because it implies that it is less important to use a particular ratio for rectangular products and packages than a ratio in the preferred range. In Study 2, the use of nine invitations with ratios extending from 1:1 to 2:1 makes it useful for an exploratory examination of whether participants had higher purchase intents within a range of ratios. We also examine each context separately because the preferred range may differ for relatively more serious versus less serious occasions.

If consumers prefer a range of ratios within the range we used in Study 2 for the two relatively serious contexts, we would expect an inverse U-shaped pattern for the purchase intent ratings. We checked for this pattern by regressing the intent ratings for the serious contexts on ratio and ratio^2 using ordinary least squares regression. The standardized beta coefficients support an inverse U shape. There is a positive coefficient for the main term and a negative one for the squared term ($\beta_{\text{ratio}} = 3.59$, $t = 17.39$, $p < .001$; $\beta_{\text{ratio squared}} = -3.74$, $t = 18.12$, $p < .001$), indicating that the relationship is not linear. Furthermore, the mean intent ratings in Figure 2 show that for both the accounting presentation and the piano concert, the ratios of 1:26–1.62 have ratings that are fairly close, and these ratings fall off monotonically as the ratios become closer to 1:1 and 2:1. The rankings data also show a similar pattern. In the relatively more serious contexts, ratios in the range of 1.26–1.5 have a higher number of first choices and a lower number of last choices than rectangles outside this range. This range does not include the Φ rectangle.

For the two relatively less serious contexts, the regression of intention ratings on ratio and ratio^2 also suggests that the relationship is not linear ($\beta_{\text{ratio}} = .59$, $t = 2.70$, $p < .01$; $\beta_{\text{ratio squared}} = -.97$, $t = -4.39$, $p < .001$). However, the mean intent ratings for both the child's birthday party and the circus trip do not exhibit an inverse U shape in Figure 2. Instead, the nonlinearity is driven by the combination of a plateau in the range of 1:1–1.5:1, with ratings decreasing for larger, more elongated ratios. Thus, there is evidence of a preferred and less-preferred range but not of an inverse U shape. Furthermore, the range of 1–1.5 has a higher number of first rankings than do higher ratios, but the data for the number of last rankings are less consistent. Thus, although we caution that these investigations are exploratory in nature, Study 2 provides preliminary evidence for a favored range rather than a particular favored rectangle for both the relatively more serious and the relatively frivolous contexts. Furthermore, these ranges and the shape of the relationships appear to differ across contexts.

Discussion

Study 2 provides additional support for H_{1a}; that is, in a consumer context, purchase intentions for rectangular consumer products depend on the ratio of the rectangle. Exploratory investigations also suggest that these intentions depend on the context of the relative seriousness of the consumer occasion for which the product is used and that preferences exist for a range of ratios rather than for a particular ratio, especially in relatively serious contexts.

Study 3: Frequently Purchased Grocery Products

Studies 1 and 2 found that the ratio of rectangular products and packages affects consumers' preferences and purchase intentions. A managerial implication of this finding is that ratio should also affect actual marketplace demand for rectangular products and packages. Whereas Studies 1 and 2 focused on purchase intentions and preferences, Study 3 examines actual marketplace demand for frequently purchased consumer goods sold in rectangular boxes. Given this focus on market demand, we test the following hypothesis:

> H_2: Marketplace demand for products sold in a rectangular box is related to the ratio of the sides of the box.

The exploratory results of Study 2 supported the notion that the relationship between purchase intention ratings and rectangular ratio was different for relatively more serious versus less serious contexts. Accordingly, we expect that the impact of ratio on marketplace demand also depends on this context. Thus, Study 3 also tests the following hypothesis:

> H_3: The direction of the relationship between marketplace demand, as measured by market share, and the ratio of the

> dimensions of rectangular boxes is different for products used in relatively more serious versus relatively less serious contexts.

Product Categories

To test H_2 and H_3, we chose four categories of products that are sold in supermarkets in rectangular boxes. To ensure comparability across categories, we chose categories in which (1) consumers see the package as a two-dimensional rectangle on the shelf but see three dimensions when inspecting the package off the shelf, (2) package materials are similar across brands (e.g., cardboard), (3) there is within-category variation in the ratio of the rectangles, and (4) there is an external criterion to measure marketplace demand. Following the method used in Study 2, we pretested four categories that met these four criteria to ensure that two of them were perceived as relatively more serious and that two were perceived as relatively less serious. A pretest ($n = 38$) showed that soaps and detergents were rated as more serious ($M = 4.15$ and $M = 4.41$, respectively, on a seven-point scale anchored by "not at all" [1] and "very serious" [7]) than cereal and cookies ($M = 3.26$ and $M = 2.93$, $ps < .05$).

Products

We identified the top eight to ten brands in terms of market shares in each category using market share data from the *Market Share Reporter* (Thomson Learning 2002, pp. 273–301, 696). The average market share in the two relatively more frivolous categories was 3.53%, whereas the average for the two relatively less frivolous categories was 9.08%. We recorded the dimensions of the height, width, and length of the box for each of these brands and calculated the following three ratios for each brand:

1. Longest dimension/median dimension (max/mid),
2. Longest dimension/shortest dimension (max/min), and
3. Median dimension/shortest dimension (mid/min).

Note that the Ratio 3 is the inverse of Ratio 1/Ratio 2. Illustrations of some of the brands considered and the variation in their shelf facings appear in Figure 3.

Results

Relationship between ratio and market share. For each of the two sets of product categories—cookies and cereals versus soaps and detergents—a separate ordinary least squares regression examined the relationship between the dependent variable of market share and the independent variables of the three ratios of the dimensions. Because market shares are percentages, we transformed them into log(share/[1 – share]). The results appear in Table 2 for the two sets of categories.

The regression results were significant for both categories ($Fs = 5.12$ and 8.19 for less and more serious categories, respectively; $ps < .05$) and explained a high degree of variation ($R^2 = .490$ and $.654$ for relatively less and more serious categories, respectively). This supports the relationship between package ratio and market share (H_2).

However, the direction of the coefficients differed across the two categories, in support of H_3. For the relatively more serious categories, the parameter estimates for max/mid and mid/min were positive ($\hat{\beta} = 63.55$ and $\hat{\beta} = 54.05$ for max/mid and mid/min, respectively; $t = 4.00$ and $t = 3.80$, respectively, $ps < .01$). In the less serious categories, the coefficient pattern was reversed: The coefficients for max/mid and mid/min were negative ($\hat{\beta} = -1.23$ and $\hat{\beta} = -1.31$ for max/mid and mid/min, respectively; $t = -2.81$ and $t = -2.82$, $ps < .01$), suggesting that for these categories, higher market shares were associated with lower ratios (i.e., closer to a square). Although the three parameter estimates for the ratios are all significant for both types of product categories, the signs for the relatively more serious categories are always opposite those for the relatively less serious categories.

Discussion

Overall, these results suggest that in the realm of frequently purchased supermarket products, the ratios of the dimensions of a product's rectangular box are related to its market share. However, because of the nature of the marketplace data, we do not claim to have identified a direction of causality; that is, H_2 and H_3 may be supported because "better" products are more appropriately aesthetically designed, but it may be that appropriately aesthetic ratios translate into demand in the marketplace. Regardless of the direction of causality, these results attest to the importance of considering product form in marketing strategy, as Bloch (1995) suggests.

General Discussion

Our studies found that the ratio of the sides of rectangular products and packages affects consumer purchase intentions and preferences and that these ratios are related to actual marketplace demand for frequently purchased consumer goods. Whereas prior work on rectangular preferences in psychology and aesthetics has typically examined rectangles as abstract entities, we examined preferences and purchase intentions in the specific domain of products and packages, which is of interest to marketing managers and product designers. Our results suggest that the proportions of a rectangular product or package can have a subtle yet important effect on consumers. On a more exploratory basis, we also found that rectangular ratio can affect consumer product perceptions and that the impact of ratio on purchase intentions and demand can depend on the relative seriousness of the context in which a product is used. Furthermore, consumers appear to favor a contiguous range of rectangles rather than a particular rectangle. However, we caution that before these results can be generalized, a broader set of rectangular products and packages must be tested.

Notably, many of the products we examined in the marketplace had ratios that were not highly favored in either our laboratory or our field studies. Thus, our findings, far from being conventional wisdom among marketing managers and product designers, may be able to inform the decisions of product designers and brand managers and add

FIGURE 3
Study 3: Examples of Products with Different Package Ratios

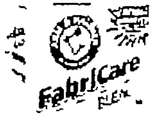
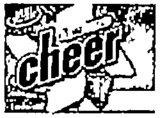
		Share (%)	Height (Inches)	Width (Inches)	Depth (Inches)	Max/Mid	Max/Min	Mid/Min	Area: Height × Width × Depth
Arm & Hammer Fabricare		3.96	10	11.19	6	1.119	1.865	1.667	671.40
Wisk		2.32	7.75	8.81	5.19	1.137	1.697	1.493	354.36
Ultra Tide with Bleach		5.49	8.63	10.88	6.31	1.261	1.724	1.368	592.47
Ultra Surf		2.06	6.75	8.75	5.19	1.296	1.686	1.301	306.53
Surf		4.94	8.63	11.25	6.25	1.304	1.800	1.381	606.80
Gain		12.94	8.13	11.25	6.31	1.384	1.783	1.288	577.13
Cheer		8.45	8	11.25	6.25	1.406	1.800	1.280	562.50
Tide		36.01	6.06	8.69	5.06	1.434	1.717	1.198	266.47

TABLE 2
Study 3: Regression Predicting Market Shares of Frequently Purchased Grocery Products (Relatively Less Serious [Cereal and Cookies] Versus More Serious [Soaps and Detergents])

	Cereal and Cookies		Soaps and Detergents	
R^2	.490		.654	
Number of observations	20		17	
Overall F	5.12 (p = .013)		8.19 (p = .003)	
Variable	**Coefficient Estimate**	**t-Value (Probability of t)**	**Coefficient Estimate**	**t-Value (Probability of t)**
Intercept	–2.19	–1.96 (.068)	–81.35	–4.04 (.0014)
Max/mid	–1.23	–2.81 (.013)	63.55	4.00 (.0015)
Max/min	.998	3.76 (.0017)	–44.11	–3.86 (.0020)
Mid/min	–1.31	–2.82 (.0122)	54.05	3.80 (.0022)

to the academic literature on product packaging (Folkes and Matta 2004; Krider, Raghubir, and Krishna 2001; Raghubir and Krishna 1999; Wansink and Van Ittersum 2003; Yang and Raghubir 2005). A brief survey of store brands showed that many private labels use a ratio that we found to be less liked. Whether this is done intentionally is an interesting area for research.[3] We hope that by demonstrating the relevance of rectangular preferences to the consumer domain, our findings will encourage researchers to explore the aesthetic appeal of other product and packaging variables and their impact on brand image, including, but not limited to, proportion, shape, color, font, relative salience, and other visual characteristics.

Our studies do not examine whether consumers' reactions to rectangular products and packages vary with culture or nationality. It is possible that consumers' environments, including the range of products and package shapes in the marketplace, the architectural norms surrounding them, the media to which they are exposed, and other cultural differences across countries and people, affect their reactions to package shapes. Cultural factors may also cause the impact of relative seriousness of context to differ across people. These issues are particularly important for international brand managers who are deciding whether to use a local or global international marketing strategy and thus are a potentially fruitful area for future research.

We examined how the shape of a rectangle affects product perceptions. Further research could examine additional product perceptions. A managerial implication of this finding is that companies may be able to change product perceptions by changing the shape of a product or package or by making shape more salient. It would also be interesting to examine whether product and package shape affects the sensory aspects (e.g., taste) or cognitive aspects (e.g., willingness to pay) of a product. Furthermore, Study 1 examined relatively serious musical concerts and the CDs for them. Additional research might examine whether the results replicate or reverse if the context is less serious, such as Cajun restaurants that encourage a fun atmosphere.

Our findings for context may help explain why prior literature on rectangular preferences has been inconclusive. As we noted previously, this research has rarely stated a context. Although participants are likely to invoke the relatively serious context of a lab experiment if provided with no other context, variations in a lab's appearance, experimenter demeanor, and instruction wording might all affect this context factor. Furthermore, of their own accord, participants often imagine in experiments that rectangles are particular objects, and there can be great variety in these objects (Ohta 1999). This could create variations in the context in which these imagined objects are used, adding an uncontrolled element to studies of preferences for rectangles.

Further research might examine whether the effects studied herein are subliminal versus conscious (and if this varies across effects) and whether they are best studied with indirect versus direct methods of elicitation of purchase intentions, preferences, or choices. It is plausible that subconscious appeals, such as aesthetic beauty, are more likely to manifest in indirect tasks, whereas more conscious and cognitive processes, such as volume estimation, may be better measured with direct tasks. Such research would be of interest given the growing interest in marketing on nonconscious processes (for a review, see Fitzsimons et al. 2002).

Our results examined how a single context, relatively more versus less serious, affects consumer reactions to rectangular packages and products. Although this context is rooted in the reasons proposed for consumers' preferences for certain rectangles, other contexts, such as whether the product is used for utilitarian or hedonic purposes, may also affect these reactions (e.g., Chandon, Wansink and Laurent [2000] show that a sales promotion's efficacy depends on the product type). Furthermore, additional research is needed to explore the construct validity of the serious/frivolous distinction to determine how it applies to a broader range of consumer products. Furthermore, we did not examine contexts that may affect how carefully or accurately consumers process information regarding rectangular shapes. This processing could be affected by color, other shapes near the rectangles, or even the rectangle's material.

---

[3] Safeway, a leading grocery chain on the West Coast of the United States, offers crisped rice bars in a box with dimensions (in inches) of 24 × 12 × 6 (shelf face ratio of box = 2), unsalted top crackers in a package with dimensions of 25.25 × 11.25 × 11 (ratio = 2.24), and sugar cones in a box with dimensions 25 × 12 × 6 (ratio = 2.08). Compare this with box dimensions of leading national brands: For example, Betty Crocker's pancake mix has dimensions of 23.5 × 16.5 × 6 (ratio =1.42), and Orville Redenbacher's popcorn has dimensions of 16 × 11.5 × 5.5 (ratio = 1.39).

Individual differences in aesthetic preferences can have an important impact on aesthetic responses to product design and shape (Bloch 1995; Holbrook 1986). Furthermore, Bloch, Brunel, and Arnold (2003) find empirical support that people differ in how important visual product aesthetics are to their "relationship with products" (p. 551); they developed a scale of the centrality of visual product aesthetics to measure this tendency, based on the dimensions of "value, acumen, and response intensity" (p. 551). Thus, an important area for further research is to examine how such individual differences may affect consumer responses to rectangular products and packages. Although gender did not have a significant impact in Study 1, it may affect the impact of rectangular ratio in other product categories. Other demographic variables also deserve study; for example, age may moderate rectangular consumer preferences if people become more serious over time.

REFERENCES

Alberti, Leon Batista (1988), *Renaissance Architectural Theoretician: On the Art of Building, in Ten Books*, Joseph Rykwert, Neil Leach, and Robert Tavernor, trans. Cambridge, MA: MIT Press. (Originally published in 1485, *De re Aedificatoria*, Florence: Nicolaus Laurentii.)

Baron, Reuben M. and David A. Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (6), 1173–82.

Benjafield, John (1976), "The 'Golden Rectangle': Some New Data," *American Journal of Psychology*, 89 (December), 737–43.

——— (1985), "A Review of Recent Research on the Golden Section," *Empirical Studies of the Arts*, 3 (2), 117–34.

——— and Keith McFarlane (1997), "Preference for Proportions as a Function of Context," *Empirical Studies of the Arts*, 15 (2), 143–51.

Bloch, Peter H. (1995), "Seeking the Ideal Form: Product Design and Consumer Response," *Journal of Marketing*, 59 (July), 16–29.

———, Frederic F. Brunel, and Todd J. Arnold (2003), "Individual Differences in the Centrality of Visual Product Aesthetics: Concept and Measurement," *Journal of Consumer Research*, 29 (March), 551–65.

Boselie, Frans (1984), "The Aesthetic Attractivity of the Golden Section," *Psychological Research*, 45 (March), 367–75.

Chandon, Pierre, Brian Wansink, and Gilles Laurent (2000), "A Congruency Framework of Sales Promotion Effectiveness," *Journal of Marketing*, 64 (October), 65–81.

DeHeer, Johan and Esther Nobbe (2000), "Is There a 'Golden' Ad? The Effect of the Spatial Layout of Ad Shape and Ad Elements on Preference," working paper, Department of Economic Psychology, Tilburg University.

Fechner, Gustav Theodor (1871), *Zur Experimentalen Aesthetik*. Leipzig, Germany: Hirzl.

——— (1876), *Vorschule der Aesthetik*. Leipzig, Germany: Breitkopf und Hartel.

——— (1997), "Various Attempts to Establish a Basic Form of Beauty: Experimental, Aesthetics, Golden Section, and Square," *Empirical Studies of the Arts*, 15 (2), 115–30. (A translation by Monika Niemann, Julia Quehl, Holger Höge, and Carl von Ossietzky of "Verschiedene Versuche, eine Grundform der Schönheit Aufzustellen. Experimentale Ästhetik, Goldner Schnitt und Quadrat," *Vorschule der Àesthetik*, Ch. 14. Leipzig: Breitkopf und Härtl.)

Fitzsimons, Gavan, J. Wesley Hutchinson, Patti Williams, Joseph W. Alba, Tanya Chartrand, Joel Huber, Frank Kardes, Geeta Menon, Priya Raghubir, J. Edward Russo, Baba Shiv, and Nader Tavassoli (2002), "Non-Conscious Influences on Consumer Choice," *Marketing Letters*, 13 (August), 269–79.

Folkes, Valerie S. and Shashi Matta (2004), "The Effect of Package Shape on Consumers' Judgments of Product Volume: Attention as a Mental Contaminant," *Journal of Consumer Research*, 31 (2), 390–401.

Ghyka, Matila (1977), *The Geometry of Art and Life*. New York: Dover Publications.

Godkewitsch, M. (1974), "The 'Golden Section': An Artifact of Stimulus Range and Measure of Preference," *American Journal of Psychology*, 87 (1–2), 269–77.

Green, Christopher D. (1995), "All That Glitters: A Review of Psychological Research on the Aesthetics of the Golden Section," *Perception*, 24 (8), 937–68.

Hekkert, Paul C., E. Peper Lieke, and Piet C.W. van Wieringen, (1994), "The Effect of Verbal Instruction and Artistic Background on the Aesthetic Judgment of Rectangles," *Empirical Studies of the Arts*, 12 (2), 185–203.

Herz-Fischler, Roger (1987), *The Mathematical History of the Golden Number*. New York: Dover Publications.

Höge, Holger (1995), "Fechner's Experimental Aesthetics and the Golden Section Hypothesis Today," *Empirical Studies of the Arts*, 13 (2), 131–48.

——— (1997), "The Golden Section Hypothesis: Its Last Funeral," *Empirical Studies of the Arts*, 15 (2), 233–55.

Holbrook, Morris B. (1986), "Aims, Concepts, and Methods for the Representation of Individual Differences in Esthetic Responses to Design Features," *Journal of Consumer Research*, 13 (December), 337–47.

Huntley, H.E. (1970), *The Divine Proportion: A Study in Mathematical Beauty*. New York: Dover Publications.

Koneni, Vladimir J. (1997), "The Vase on the Mantelpiece: The Golden Section in Context," *Empirical Studies of the Arts*, 15 (2), 177–207.

Krider, Robert E., Priya Raghubir, and Aradhna Krishna (2001), "Pizzas: π or Square? Psychophysical Biases in Area Comparisons," *Marketing Science*, 20 (4), 405–425.

Lalo, C. (1908), *L'Esthetique Experimentale Contemporaine*. Paris: Alcan.

Lawlor, Robert (1982), *Sacred Geometry: Philosophy and Practice*. London: Thames & Hudson.

Livio, Mario (2002), *The Golden Ratio: The Story of Phi, the World's Most Astonishing Number*. New York: Broadway Books/Random House.

McManus, I.C. (1980), "The Aesthetics of Simple Figures," *British Journal of Psychology*, 71 (November), 505–524.

——— and P. Weatherby (1997), "The Golden Section and the Aesthetics of Form and Composition: A Cognitive Model," *Empirical Studies of the Arts*, 15 (2), 209–232.

Miller, Cyndee (1994), "The Shape of Things: Beverages Sport New Packaging to Stand Out from Crowd," *Marketing News*, 28 (17), 1–2.

Ohta, Hirohiko (1999), "Preferences in Quadrangles Reconsidered," *Perception*, 28 (4), 505–517.

Piehl, Jochen (1976), "The Golden Section: An Artifact of Stimulus Range and Demand Characteristics," *Perceptual and Motor Skills*, 43 (1), 47–50.

——— (1978), "The Golden Section: The 'True' Ratio?" *Perceptual and Motor Skills*, 46 (1), 831–34.

Plug, C. (1976), "The Psychophysics of Form: Scaling the Perceived Shape of Plane Figures," *South African Journal of Psychology*, 6, 9–17.

——— (1980), "The Golden Section Hypothesis," *American Journal of Psychology*, 93 (3), 467–87.

Prince, Greg W. (1994), "The Contour: A Packaging Vision Seen Through Coke-Bottle Lenses," *Beverage World*, (May 31), 1, 6.

Raghubir, Priya and A. Krishna (1999), "Vital Dimensions in Volume Perception: Can the Eye Fool the Stomach?" *Journal of Marketing Research*, 36 (August), 313–26.

Sherwood, Mark (1999), "Winning the Shelf Wars," *Global Cosmetic Industry*, 164 (March), 64–67.

Svensson, Leif T. (1977), "Note on the Golden Section," *Scandinavian Journal of Psychology*, 18 (1), 79–80.

Thomson Learning (2002), *Market Share Reporter*. Detroit: Thomson Learning.

Thorndike, Edward Lee (1917), "Individual Differences in Judgements of the Beauty of Simple Forms," *Psychological Review*, 24, 147–53.

Wansink, Brian and Koert van Ittersum (2003), "Bottoms Up! The Influence of Elongation on Pouring and Consumption Volume," *Journal of Consumer Research*, 30 (December), 455–63.

Yang, Sha and Priya Raghubir (2005), "Can Bottles Speak Volumes? The Effect of Package Shape on How Much to Buy," *Journal of Retailing*, 81 (4), 269–81.

Eyal Biyalogorsky, William Boulding, & Richard Staelin

Stuck in the Past: Why Managers Persist with New Product Failures

In this research, the authors examine the phenomenon of escalation bias in the context of managing new product introductions. In particular, they identify three general paths—Decision Involvement Inertia, Decision Involvement Distortion, and Belief Inertia Distortion—that can lead managers to escalate their commitments. The authors test the relative strength of these paths in driving observed escalation behavior. The results show that involvement with the initial decision, a key construct in numerous explanations for escalation behavior (e.g., agency theory, self-justification), is not a necessary condition to induce commitment to a losing course of action (i.e., escalation bias). Rather, the authors find that the driving force behind escalation behavior is improper use of initial positive beliefs in the face of negative new information. This insight has implications for the groundwork necessary for organizations to design systems, policies, and procedures to help them avoid the trap of escalation bias that is often associated with major strategic decisions.

Marketing managers are often asked to make major decisions, even though the information available to them at the time of the decision is incomplete and uncertain. Typical examples of such decision situations are launching a new product, initiating an everyday-low-pricing or a "no-promotions" strategy, opening a new channel of distribution, and initiating a major new media campaign. Given that the decision environment is often highly uncertain and the financial stakes large, it is not surprising that firms often revisit such decisions after they obtain new information to determine whether the firm should continue with, modify, or terminate the initially chosen course of action.

It is well documented that when managers who are publicly committed to a course of action are asked to reevaluate that action, they tend to remain committed to it even when the new information indicates that the action should be terminated. See, for example, studies of the Vietnam War and Desert Storm (Lipshitz 1991), Expo 86 (Ross and Staw 1986), the Apollo moon missions (Mitroff 1974), the Campeau–Federated merger, the coffee wars between Philip Morris and Procter & Gamble (Bazerman and Neale1992), the National Basketball Association draft (Staw and Hoang 1995), information technology projects (Keil 1995; Keil, Mann, and Rai 2000), and a plethora of experimental investigations (Arkes and Blumer 1985; Boulding, Morgan, and Staelin 1997; Brockner and Rubin 1985; Staw 1976). This stylized phenomenon, often referred to as "escalation of commitment," can be disastrous for firms, especially in today's hypercompetitive markets, which require fast and accurate reactions and adaptability on the part of companies.

Decisions about new products seem particularly susceptible to the problem of escalation of commitment. Schmidt and Calantone (1998) show that innovative projects tend to generate a high level of commitment from managers, often leading to ongoing investment in failing projects. Industry practitioners echo this sentiment. The director of the New Products Showcase and Learning Center stated, "It sometimes takes more courage to kill a product that's going nowhere than to sustain it" (Lukas 1998, p. 44). Still, killing a new product can be the best course of action, as suggested by comments from a former chief executive officer of Eastman Kodak when speaking to participants in a senior executive education program. He indicated that he considered a new product team a "success" when members appropriately decided to *stop* investing in a new product idea and labeled the project "the world's fastest failure."

We use the new product introduction setting to explore the reasons managers tend to stay with losing courses of action and thus forgo the best course of action. We begin by developing a model of how managers combine specific information about the potential product and their prior business experience to decide whether to launch a new product and, if they choose to launch the product, how they decide whether to continue with the venture after receiving negative new information. We then estimate the parameters of this model using data that we obtain from an experiment in which managers are asked to make an initial product-launch decision and a subsequent reevaluation decision. Using the resulting estimates, we provide empirical support for and against numerous previous explanations of escalation of commitment. In turn, this enables us to address the important question of how firms can reduce managers' tendency to stay with losing courses of action.

Paths to Escalation

To make the preceding discussion more specific, imagine that a manager is provided with a detailed analysis of the potential viability of a new product. This analysis contains

---

Eyal Biyalogorsky is Associate Professor of Management, Graduate School of Management, University of California, Davis (e-mail: eyalog@ucdavis.edu). William Boulding is Professor of Marketing (e-mail: bb1@duke.edu), and Richard Staelin is Edward and Rose Donnell Professor of Business Administration (e-mail: rstaelin@duke.edu), Fuqua School of Business, Duke University.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

information about specific aspects of the market—the distribution of possible initial market shares, industry growth rates, likely responses by competitors, and summary financial data, such as the anticipated net present values (NPVs) of the venture. The manager uses all of this information to form beliefs about what is likely to happen and, on the basis of these beliefs, comes to an initial decision about whether to launch the new product.

Assume that the manager in our example decides to introduce the product. Moreover, assume that this manager is assigned to lead the product launch so that he or she is publicly connected to the initial launch decision. Two years later, the firm decides to reevaluate the viability of the product. The analyst again provides the manager with information, this time based on new data collected since the launch decision. Among other things, the new information contains updated figures for realized market share, new estimates of industry growth rates, and newly anticipated competitive responses. This information is also summarized in terms of the anticipated distribution of NPVs, assuming that the firm continues with the product. In addition, the analyst provides information about the expected NPV of exiting the market and using the assets associated with the venture in another application. Finally, assume that the "normative" decision (i.e., the decision based on the expected outcomes associated with sticking with the product relative to the expected alternative use of funds) is to pull the plug on the product and reinvest the remaining assets. When the manager is asked either to stay the course or to reinvest the assets in the new application, he or she decides to stay the course, demonstrating escalation of commitment.

Why might escalation bias occur? The extant literature can be organized around three general paths that lead to this occurrence. The first path implicitly or explicitly assumes that because the decision maker was publicly involved with the first decision, this involvement directly influences the second decision. Returning to our example, this means that the manager continues with the product launch despite being fully aware of the negative new information because of his or her involvement with the initial decision. This may happen because the manager fears "losing face" if the initial decision is changed (Brockner, Rubin, and Lang 1981) or because of the effort required to justify the change (Fox and Staw 1979). Another possibility is that the manager uses the NPV assessment associated with the first decision as a reference point and consequently perceives a lower assessment as a loss. Prospect theory posits that people tend to become more risk seeking in the domain of losses (Kahneman and Tversky 1979). Framing the issue as a loss may lead the manager to accept more risks than normal and therefore stick with the initial launch decision (Arkes and Blumer 1985; Bazerman 1984; Whyte 1986). In all of these explanations, the manager's involvement with the initial decision leads to inertia and reluctance to change, even though the negative new information may be correctly perceived. Thus, we refer to this path as "Decision Involvement Inertia."

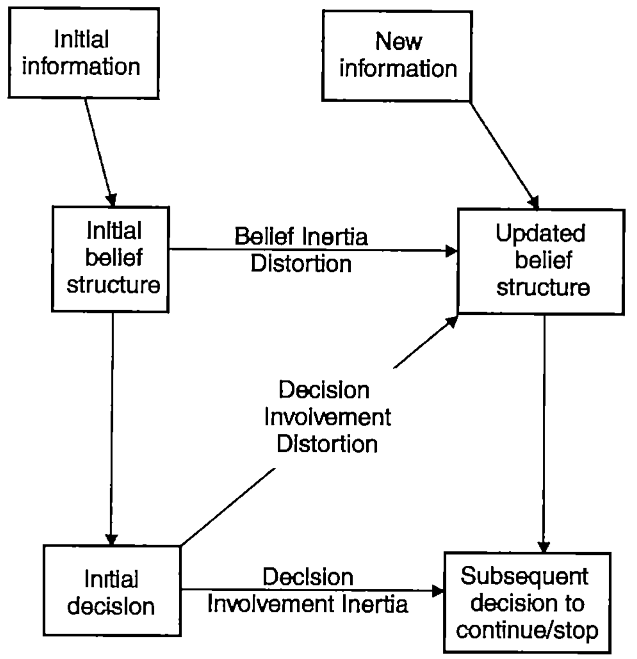
The manager's involvement with the initial decision again plays a critical role for the second major escalation path. Here, involvement with the first decision does not affect the second decision directly; instead, it improperly influences the belief structure underlying the second decision (Bacharach, Bamberger, and Mundell 1995; Conlon and Garland 1993; Festinger 1957; Samuelson and Zeckhauser 1988; Silver and Mitchell 1990; Wicklund and Brehm 1976). The manager, who was involved with the first decision, misinterprets the new information so that it is in concert with his or her initial decision, resulting in biased beliefs that, in turn, lead to the decision to stay the course. Therefore, we call this path "Decision Involvement Distortion" because the manager's involvement with the initial decision leads to a distortion in the perception of the new information.

The third path also concerns biased perceptions of the new information. However, in this case, the bias comes from the person distorting and/or improperly weighting the new information to conform to his or her initial beliefs, not from his or her involvement with the initial decision. Thus, returning to our example, this mechanism suggests that it is not necessary for the manager to have made the initial product introduction decision to exhibit commitment escalation subsequently. Rather, it is necessary only that the manager has formed positive beliefs about the viability of the initial product launch, even with no public involvement in the launch decision, to induce escalation. In other words, this bias does not occur because the manager was involved in the initial decision but rather because he or she does not sufficiently adjust prior beliefs and thus chooses to continue the original course of action. Note that the mechanism for these biased beliefs is similar to that associated with a confirmatory bias or sticky prior beliefs (Bolton 2003; Jonas et al. 2001; Klayman 1995; Schwenk 1986). Thus, we refer to this path of bias as "Belief Inertia Distortion."

In summary, the literature on escalation can be organized around three basic mechanisms (paths) that lead to escalation of commitment. We summarize these mechanisms in Figure 1. Two of the mechanisms rely on the maker of the second decision being involved in the first decision. The third relies only on the decision maker initially having positive beliefs about the factors that led to the first decision. All three mechanisms can cause managers to exhibit escalation behavior, but they differ with respect to the process that causes the behavior. Moreover, they differ in terms of their implications regarding how to prevent escalation of commitment.

Determining which paths are operative is not an easy task. All the paths involve complex relationships between second-stage outcomes and first-stage explanatory variables. It is not enough to demonstrate simple correlations between first- and second-stage variables. For example, to test Decision Involvement Inertia, it is necessary to show that involvement with the initial decision influences the second decision, after we control for the effects of the updated second-stage belief structure. Likewise, testing Decision Involvement Distortion requires us to show that the manager's updated belief structure affects the continue/stop decision and then that involvement with the initial go/no-go decision affects the manager's updated belief structure, after we control for the possible normative and nonnormative effects of prior beliefs on updated ones. Finally, to test Belief Inertia Distortion, we must show that the manager's

FIGURE 1
Paths Leading to Escalation


Initial information
New information
Initial belief structure
Belief Inertia Distortion
Updated belief structure
Decision Involvement Distortion
Initial decision
Decision Involvement Inertia
Subsequent decision to continue/stop

updated belief structure affects the continue/stop decision and that the manager's prior beliefs have a nonnormative effect on his or her updated beliefs beyond any predicted normative effect on the updated beliefs, independent of any involvement with the initial decision.

The preceding discussion makes it clear that if we want to disentangle the potential influences of the first decision from the influence of prior beliefs, we must explicitly model and measure the belief-updating process. In Heath's (1995) review of the literature, he expresses concern that most escalation studies fail to control for initial beliefs. Consequently, we take extra care to address this issue.

Given the complexities associated with disentangling our effects of interest, we chose to conduct an experiment in which we manipulate the manager's involvement with the initial decision, the level of the manager's initial (prior) beliefs, and the degree of (lack of) market success subsequent to the initial product launch. We then measure initial and updated beliefs and observe the second decision. This enables us to estimate the parameters of our belief-updating model and, on the basis of those parameters, to assess the impact of each of the three potential paths on escalation of commitment. Formally, we test the following three hypotheses:

> H_1 (Decision Involvement Inertia): When the decision maker's updated belief structure is held fixed, being highly involved with the decision to choose an initial course of action has a direct, positive effect on the next decision, thus producing escalation behavior.
>
> H_2 (Decision Involvement Distortion): Being highly involved with the decision to choose an initial course of action biases the decision maker's updated belief structure in the direction of supporting the initial course of action. In turn, this biased updating process increases the likelihood that the decision maker will stay with the initial course of action.
>
> H_3 (Belief Inertia Distortion): Independent of any involvement with the initial decision, a decision maker's initial positive belief biases his or her updated belief structure in the direction of the initial belief structure, thus producing escalation behavior.

Note that these three hypotheses are not mutually exclusive and therefore can operate simultaneously. Furthermore, any empirical tests of these hypotheses must distinguish between normative and nonnormative updating of beliefs. Consequently, in the next section, we describe a model for belief updating and, in the process, describe how we can distinguish between normative and nonnormative updating effects in the model.

Belief-Updating Model

Our model of the belief-updating process acknowledges that a marketing manager normally receives different and, perhaps, conflicting pieces of information about many aspects of the environment. We assume that the manager integrates this information into an overall assessment that guides a decision. Therefore, we model not only the manager's overall assessment and updating process but also the process of updating beliefs about specific underlying aspects of the environment that lead to the overall assessment.

We begin with the conceptualization of managerial decision making that Boulding and colleagues (1994) propose. Their model assumes that all beliefs can be represented by probability distributions in which the mean of the distribution is the person's belief about the most likely occurrence and the variance of the distribution captures the person's uncertainty about the particular belief. Boulding and colleagues posit two different types of underlying beliefs and one summary belief. The underlying beliefs refer to the specific levels of the market factors that influence performance of the product (also referred to as "what" knowledge) and beliefs about the market structure (i.e., how each market-factor belief is weighted; also referred to as "how" knowledge). This conceptualization has the manager combine the two components (factor levels and weights) to form "summary" knowledge, that is, a summary belief or overall evaluation. For example, for a new product decision, market-factor beliefs would include all the factors that are believed to have an influence on the project's overall success. These market factors could include items such as the obtained market share, the price of the new product offering, market growth, competitors' responses, and the cost of capital. The manager then uses knowledge of the relevant market structure to establish weights for how each market-factor belief affects the project's overall success, which is captured by the manager's summary belief. Boulding and colleagues assume that it is this summary belief that ultimately determines the manager's decision.

We formalize this process as follows: Let SB_{it} be the mean of manager i's summary belief about the course of action at time t, and let FB_{ikt} be the mean of the manager's belief about the level of the kth market factor that affects overall market performance. Let $g_{it}(\cdot)$ be the manager's mental mapping of how these market factors combine to result in overall performance:

$$SB_{it} = g_{it}(FB_{ikt}). \tag{1}$$

We assume that the manager's mental mapping, $g_{it}(\cdot)$, can be closely approximated by a Fishbein-style linear formulation (Ajzen and Fishbein 1977). Thus, in Equation 2, we replace the function $g_{it}(\cdot)$ with a vector of importance weights (w_{ikt}) that captures the manager's perception of the importance of each market factor:

$$SB_{it} = \sum_{k} w_{ikt} FB_{ikt}. \tag{2}$$

Conceptually, the assumption of a linear mapping is supported by numerous studies that have shown that a linear model provides a robust, parsimonious, and paramorphic representation of how people combine multiple stimuli when forming an overall evaluation (e.g., Ajzen and Fishbein 1977; Bettman 1979; Einhorn 1970; Johnson and Meyer 1984; Keeney and Raiffa 1976). Equally important, the data-generating process we used to create the information presented to participants was linear. Thus, participants who correctly perceived the relationships in the data should have used a linear mapping. Empirically, we tested several nonlinear specifications for the participants in the study and found no indication that a nonlinear specification was necessary or appropriate.

Consider now SB_{i2}, which represents the manager's summary belief at the continue/stop decision point. Equation 2 suggests that there are three potential sources of bias of this summary belief: (1) The manager's market-factor beliefs, FB_{ik2}, could differ from normatively correct levels; (2) the manager's importance weights, w_{ik2}, could be distorted compared with normatively correct weights; and (3) when forming a summary belief, the manager could inject extraneous, irrelevant information that Equation 2 does not capture. For example, if the manager had an argument with his or her children over breakfast, an ensuing bad mood could affect his or her beliefs. Theoretically, we can capture this effect in Equation 2 by augmenting the market-factor-beliefs vector to include the manager's mood. As researchers, we have access to and can measure only a limited set of possible factors, and thus any estimation must try to control for unobserved factors.

Next, we describe how to determine whether any of these sources of bias affect the manager's updated beliefs at the continue/stop decision stage. In doing so, we are interested not in just any difference between the manager's beliefs and the normatively correct beliefs but rather in the possible systematic biasing effects of the manager's first-stage (go/no-go) beliefs (SB_{i1}, FB_{ik1}, w_{ik1}) and the initial decision on his or her second-stage beliefs.

Updating Market-Factor Beliefs

We begin by modeling how a manager updates the mean of his or her market-factor beliefs (see Equation 3). The model includes three terms: (1) the effect of any relevant new information that the manager receives about market factors (NI_k), (2) the effect of the manager's prior market-factor belief (FB_{ik1}), and (3) the effect of the manager's prior summary belief (SB_{i1}).

$$FB_{ik2} = \alpha_{1k} NI_k + \alpha_{2k} FB_{ik1} + \alpha_{3k} SB_{i1}. \tag{3}$$

Note that we do not explicitly model the effects of the initial decision on the updating process. Rather, we hypothesize that if the person's involvement with the initial decision affects the updating process, the coefficients in Equation 3 will systematically vary depending on the level of involvement. In other words, involvement interacts with the right-hand side of Equation 3. In addition, note that if we assume that the manager's beliefs can be described by a normal distribution, the first two terms are consistent with normative updating, where the coefficients α_{1k} and α_{2k} reflect the relative reliabilities of the new information and the initial belief, respectively. Conversely, the third term represents the hypothesized nonnormative effect of the initial summary belief on the updating of market-factor beliefs. The rationale behind this last term is the well-documented phenomenon that people tend to perceive new information as compatible with their prior beliefs, exhibiting a confirmatory bias (Hoch and Ha 1986). A significant, positive coefficient on the initial summary belief is evidence of biased updating of the market-factor belief.[1]

Updating Importance Weights

Normatively, importance weights should change only if the manager receives new information that specifically indicates that the model of how market factors interact has changed. However, we conjecture that new information that pertains only to the levels of market factors may still induce a change in importance weights. Specifically, if the manager's assessment of a market-factor level increases (decreases), we suggest that there will be an associated increase (decrease) in the importance that he or she accords to that market factor. Such a pattern would reflect "elastic" importance weights (Boulding, Kalra, and Staelin 1999; Hsee 1995). Specifically, the weights would increase in value when new information supports a positive initial belief and decrease in value when new information contradicts the positive initial belief. In Equation 4, the third term captures this reasoning; if the difference between the manager's new market-factor belief and prior market-factor belief is positive (i.e., the new information is perceived as more positive), the importance weight for this factor increases. If the difference is negative, the reverse occurs. As in Equation 3, the first two terms in Equation 4 represent the expected normative effects.

$$w_{ik2} = \beta_{1k} NI_k + \beta_{2k} w_{ik1} + \beta_{3k}(FB_{ik2} - FB_{ik1}). \tag{4}$$

---

[1] Failing to find a significant effect on the initial summary belief does not mean that the updating is normative, because the other two coefficients may still be different from their normative values. Here, however, we restrict our attention to the more stringent test of the significance of the prior summary-belief coefficient.

We estimate a slightly different version of Equation 4. It will become clearer when we describe the experiment that we do not manipulate or provide any new information that indicates a change in the market structure (the importance weights). Thus, β_{1k} should be zero, and we estimate only the last two parameters in Equation 4:

$$(5) \qquad w_{ik2} = \beta_{2k}w_{ik1} + \beta_{3k}(FB_{ik2} - FB_{ik1}).$$

In this "special case updating" given in Equation 5, $\beta_{2k} = 1$ and $\beta_{3k} = 0$ would be consistent with normative updating, and deviations from these parameter values would reflect biased updating.

Updating the Summary Belief

If the manager's market-factor beliefs and/or importance weights at the time of the continue/stop decision are biased, then according to Equation 2, the summary belief that the manager constructs from those underlying beliefs will be biased as well. However, it is possible that the manager's updated summary belief is not constructed in this manner. Rather, the manager might update the former summary belief on the basis of some macrolevel evaluation of the veracity of the new information. Such a model is consistent with normative Bayes-rule updating, but at a higher level of aggregation. Because the two possible updating models are based on the same information, it follows that the two should agree in a well-calibrated model. On the basis of this principle, we can test for the existence of bias at the summary-belief level. Consider a general model that nests our two alternative models of summary-belief updating:

$$(6) \qquad SB_{i2} = \gamma_1 NI + \gamma_2 SB_{i1} + \gamma_3 \Sigma w_{ik2} FB_{ik2}.$$

If the mean of the summary belief is updated by means of Bayes' rule with the prior summary belief and the new information, the γ_3 term in Equation 6 should provide no additional explanatory power. Conversely, if the updated belief is constructed on the basis of the updated underlying factor beliefs, the first two terms in Equation 6 should provide no additional explanatory power. If all three terms simultaneously provide explanatory power for the updated summary-belief measures, bias is entering the summary belief.

In summary, estimation of Equations 3, 5, and 6 enables us to test directly for the existence of bias in the updating process and, therefore, the mechanisms underlying escalation behavior (H_1–H_3).

Method

Before describing our study, we briefly review the requirements needed to test adequately for the occurrence of the three proposed mechanisms underlying escalation behavior. First, as Northcraft and Wolfe (1984, p. 227) point out, we need to "examine decision making situations in which commitment of further resources is explicitly economically inadvisable." Heath (1995, p. 52) also mentions this need and further claims that "previous demonstrations of 'escalating commitment' with explicit investments did not provide relevant information about marginal costs and benefits." To address these concerns, we provide the decision makers with explicit information about the marginal costs and benefits of both continuing with and discontinuing the course of action. Furthermore, the information shows that to continue is economically inadvisable. Second, because involvement with the initial decision and initial beliefs are key constructs in the three paths, it is necessary to manipulate both involvement with the first decision and initial beliefs. This is not straightforward, because measuring beliefs requires the person to be somewhat involved with the initial decision. However, using several different manipulation checks, we show that even though all our participants are somewhat involved with the initial decision, our involvement manipulation leads to significant differences in their involvement level. Finally, there is the issue of disentangling normative belief updating from biased updating. Here, we use our theoretical model to partition the two effects.

More specifically, we conducted an experiment in which we asked participants to take the role of management in a new product introduction and reevaluation situation. The study involved manipulating three variables: the degree to which participants were involved with the initial decision (high involvement versus low involvement); the extent to which the initial information provided a positive initial indication about future profits and, thus, positive NPV (very positive versus positive); and the degree to which subsequent information was nonconfirmatory (very negative versus negative). Together, these three manipulations enable us to tease apart the various effects related to H_1–H_3.

We used a new product setting similar to our initial example that appears in a previous study (Boulding, Morgan, and Staelin 1997) in which senior-level managers exhibited a marked tendency to continue with the product. Specifically, between 45% and 100% of the managers continued with the product, depending on the particular condition. In contrast, when Boulding, Morgan, and Staelin (1997) gave the financial information associated with the reevaluation decision to a group of 20 academic experts, all chose to discontinue the project. This latter finding provides an important normative benchmark, suggesting that continued commitment to the project in our setting is indeed evidence of escalation bias.

Participants

We obtained responses from 142 participants who completed the experiment. Approximately two-thirds of these participants were MBA students from four different full-time programs, two on the East Coast and two on the West Coast. The remaining participants were midlevel managers who were enrolled in an executive MBA program at one of the East Coast schools. All participants had completed at least one full semester of studies that included instruction in decision making. Because our sample used MBA students, it is logical to ask if this sample can be generalized to the population of senior-level managers who normally make major investment decisions. To address this issue, we replicated the control condition used in Boulding, Morgan, and Staelin's (1997) study, which used senior-level managers as participants, with a sample of our participants. In our study, 60% of the participants (12 of 20) in the control condition discontinued the product, whereas 66% (27 of 41) discontinued it in Boulding, Morgan, and Staelin's study.[2] The difference in stopping rates between the two studies is not significant ($t = .45, p > .30$). This point of comparison suggests that there is no substantial difference in the way participants in the two studies responded to the information provided, increasing our confidence that we can generalize our findings to the population of senior-level managers sampled in Boulding, Morgan, and Staelin's study.

The Case

Participants were asked to complete the Quality Valve Company case.[3] The first two pages of the case describe the new product, the history of the market, the company, the primary competitor, and the investment associated with an improved valve targeted at the large-truck emissions-control market. The specific initial decision problem was whether the organization should invest $2.5 million in new machinery and launch the improved valve.

As part of the case description, participants learned that the financial success of the new product venture depended on three key uncertainties: the share of the market the new product garners after launch, the industry's growth rate, and the competitive-entry response of the major competitor to the launch of the new product (which was captured by the timing and effectiveness of an innovative competitive offering in the future).

Participants initially received one of two different analyst's reports (positive or very positive) that included projections of all the key market factors along with the distribution of estimated NPVs, assuming that the product is launched. They also were given the expected NPV and could easily determine from the NPV distribution the probability that the project would yield a positive NPV. In addition, the analyst's report provided estimates of the relative importance of each of the key uncertainties in determining the product's financial success; these estimates are 70% for market share, 22% for competitive-entry response, and 8% for market growth. After reading the case description and the analyst's report, participants made a launch/no-launch decision (high-involvement condition only) and described their beliefs about the expected levels of relevant market factors and their mean overall summary beliefs.

Next, participants in the high-involvement group who suggested launching the product and all the participants in the low-involvement group were given the second part of the case. Specifically, they received negative new information about the product's performance during its first two years on the market. The information included realized values for the three key market uncertainties and was presented in a balanced fashion. Thus, the realized market share was significantly less than anticipated, the realized market growth rate was greater than anticipated, and the information about the major competitor's entry response was neutral in that no new information was given on either the timing or the effectiveness of its future offering. In addition, participants were told that an uncorrectable production problem explained the lower-than-expected market share but that the sales force and research-and-development groups were very happy about the product introduction. Nonetheless, because market share was the most important factor in determining profitability and the production problem was not correctable, on balance, the information was negative. Moreover, the analyst's revised (normatively updated) financial report predicted an expected NPV from future cash flows that was much lower than initially projected. More important, this NPV forecast was substantially lower than the guaranteed cash value associated with the alternative of dropping the product and selling the equipment. Thus, as Heath (1995) and Camerer and Weber (1999) stress, the marginal costs and benefits of the alternative courses of action were made explicit to participants before they made the second decision. Specifically, the normative decision for the organization based solely on the new financial information and the concept of expected values was to withdraw the product and sell the equipment. This choice was given to participants. After receiving the new information, participants were asked to select one of two alternatives: continue with the product or withdraw it and sell the equipment.

Experimental Manipulations

Initial information. The initial information provided to participants was either positive or very positive. In the very positive condition, the expected market share following the launch was higher, as was the analyst's assessment of the expected NPV from the product. In both conditions, the information was sufficiently positive to result in the vast majority of participants choosing to launch the product.

New information. The new information provided to participants was either negative or very negative. In the very negative condition, participants were given a realized market share after the launch that was lower than the market share given to participants in the negative condition. In addition, the analyst's assessment of the expected NPV from the product was lower in the very negative condition. In both conditions, the information about realized market share, the uncorrectable production problem, and the analyst's revised NPV assessment declined sufficiently compared with the guaranteed cash-out value of stopping the project to make stopping the project the normatively correct decision.

Involvement manipulation. The theoretical explanations underlying the two decision involvement paths suggest that the influence of the initial decision on the subsequent continue/stop decision is the result of the decision maker's personal involvement in the initial decision (see, e.g., Staw and Ross 1987). The literature on escalation bias has manipulated involvement with a decision by

---

[2] Although this control cell replicates Boulding, Morgan, and Staelin's (1997) study, the levels of information that they provide differ slightly from those used in this study. Consequently, these data cannot be used for further comparisons with other data collected in this research.

[3] The case is available on request.

1. Manipulating the level of participation the second-stage decision maker had with the initial decision by having him or her participate (or not) in making the initial decision (Barton, Duchon, and Dunegan 1989; Staw 1976), making him or her primarily responsible for the decision or responsible only for making a recommendation to be reviewed by others (McAllister, Mitchell, and Beach 1979), and making him or her the company president or someone with no connection to the company (Arkes and Blumer 1985);
2. Manipulating the personal stake the decision maker had in the success of the initial course of action (Arkes and Blumer 1985) by, for example, implying that there were or were not job security implications (Fox and Staw 1979); and
3. Manipulating how much the decision maker expected to need to justify the decision to others (Bobocel and Meyer 1994; Simonson and Staw 1992).

We used the first two methods to manipulate involvement with the initial decision. Specifically, high-involvement participants were put in the position of making the initial decision to launch or not to launch the product after receiving the previously discussed information. These participants were instructed that if they launched the product and it succeeded, they would likely be on the fast track to a senior management position. They were told that choosing not to launch the product would not hurt their careers but that they would lose this particular opportunity to be on the fast track. Those who chose to launch the product were subsequently told that they were appointed as the product's brand manager. Consequently, they had overall responsibility for everything that happened, beginning with the decision to launch the product.

Conversely, low-involvement participants were told that the boss, who had a highly centralized decision-making style, would make the launch decision. However, the boss wanted to know the manager's beliefs about the key uncertainties associated with the launch. Furthermore, participants in this condition were told that the boss wanted unbiased opinions, and thus all evaluations about these key factors were to be provided anonymously. After they provided the requested belief information, these low-involvement participants were told that during the two years the product was on the market they were assigned to an entirely different project and that they were being called back to help make the reevaluation decision. Thus, the high- and low-involvement conditions differed in terms of (1) the level of the decision maker's involvement in the initial launch decision (similar to the method that McAllister, Mitchell, and Beach [1979] used), (2) the level of the decision maker's involvement with the product's performance after launch (Arkes and Blumer 1985), and (3) the potential career implications of the success of the initial decision for the decision maker (Arkes and Blumer 1985).

To ensure that we successfully manipulated involvement with the initial decision, we conducted two independent manipulation checks. We provide details of these two checks in the Appendix. In summary, for all manipulation checks, there were significant differences in the expected directions between the two groups with respect to their involvement with the initial decision. Still, we acknowledge that the low-involvement situation required the decision maker to think hard about the different aspects of the decision and, in this way, differed from some prior studies in which the decision maker was not provided any information about the first decision. However, as Heath (1995) points out, it is important to capture the participants' beliefs before the first decision is made. This required the participants to know about the decision and form beliefs about the elements of the decision that were relevant at the time. Consequently, all participants were at least somewhat "involved" with the decision. In the "Results" section, we discuss a third "manipulation check" that is based on the participants' actual actions. To foreshadow these results, again we found that the two groups differed, but this time with respect to their decision behavior. This provides additional support for our contention that there are significant differences in the two involvement groups.

Design

We used an incomplete 2 × 2 × 2 design for the study. The low-involvement condition was fixed at the positive level of the initial information (as opposed to very positive), but it was crossed with both levels of new information. Thus, the design included only six of the eight possible cells. We used this design for parsimony because there are no theoretically meaningful predictions about the interaction of involvement with the initial decision and the level of new information.

Measures

Our theory testing revolved around three different measured constructs: market-factor beliefs, importance weights, and summary beliefs. After each decision, we took two summary-belief measures. They were the manager's assessment of expected NPV and the probability of success (the probability that NPV would be positive). We combined these measures after normalizing each measure over the sample to have a mean of zero and a standard deviation of one. We then summed the two measures for each participant to form the summary-belief measure used in our analysis.

We took measures for each of the three market-factor beliefs after both decisions. We measured the market growth rate and market-share beliefs using a nine-point scale anchored between "very negative" and "very positive" (the midpoint was designated as "neutral"). Our use of affective evaluations rather than actual expectations ensured that all participants were on the same "goodness" metric. In this way, a higher rating reflects a higher evaluation even if, for example, a particular manager believes that a 20% market share is a great success whereas another considers a 40% market share a relative failure. For the market-factor belief associated with the major competitor's entry response, we used two measures: the likelihood of competitive launch and the effect of a competitive launch. We measured both on nine-point scales, normalized them, and combined them into a single variable that, as with our summary-belief measure, captured both the probability of the event occurring and the value of the event.

Finally, we measured importance weights for the components relevant to the second decision by asking participants to specify the relative importance of the three market factors in determining the project's financial success. In particular, participants divided 100 points between the three factors.

Results

Because we are interested in situations that involve two decisions, we discarded data for the 7 high-involvement participants who chose not to launch and for the 5 low-involvement participants who, given their beliefs, presumably would have chosen not to launch.[4] Of the remaining 130 participants, 68 (52%) decided to stay with the losing course of action, exhibiting behavior that is compatible with an escalation bias. Given this evidence of nonnormative behavior, we next report findings that bear on the strength of the three nonnormative paths based on our framework.

Decision Involvement Inertia Results

The Decision Involvement Inertia path implies that involvement should directly influence the second decision. To test this, we heeded the warnings of Heath (1995) and controlled for summary beliefs. Table 1 presents the results of our logistic regression for three different models. In Model 1, the only independent variables are two of our experimental manipulations—initial information and new information—and our measure of the participant's summary beliefs. Note that summary beliefs are highly significant and completely mediate any effects of the initial and new information. In Model 2, we add a variable that captures the experimental manipulation of involvement with the initial decision. As is immediately evident, involvement does not reach traditional levels of significance ($p = .12$). Moreover, there are no significant differences between Model 2 and Model 1 in the log-likelihood (122.135 versus 124.609) or in the predicative accuracy (84.3% versus 82.6%). Thus, we find little statistical justification for the Decision Involvement Inertia path.

A question might arise as to why our involvement measure did not affect the second decision, given that numerous previous studies have shown an effect. One explanation is that our involvement manipulation did not provide enough difference between the two groups in terms of their perceptions of being involved with the first decision. However, when we ran a model without the summary-belief measure but with the involvement measure (Model 3), we found that involvement was positively associated with continuing with the initial course of action and that this effect was significant at the .05 level. This result has three major implications: First, it provides one more piece of evidence that our involvement manipulation worked. Second, it reproduces many previous findings on involvement in which researchers did not control for summary beliefs. Third, a comparison of this result with the insignificant involvement effect in Model 2 shows that even though involvement may be associated with escalation of commitment, being psychologically and/or socially involved with the initial decision is not a necessary condition for escalation to occur. Instead, it seems that escalation bias enters the decision through the updated belief structure. Consequently, we turn our attention to the two paths that pertain to distortion of information and, thus, biased updating.

Decision Involvement Distortion Results

Before we present the Decision Involvement Distortion results, note that we tested both distortion paths using some variables that we measured instead of manipulated. This required us to control for background factors, such as participants' individual differences, that could be correlated with both our independent variables and our dependent measure of interest (and thus produce misleading estimates). We did this using an instrumental variable (IV) procedure, in which the selected instruments are known to be free from any individual background factors. The IV procedure replaces the actual value of a measure with its pre-

TABLE 1
Results of Logistic Regression on the Continue/Stop Decision (N = 126)

Variable	Model 1	Model 2	Model 3
Intercept	–4.94*	–5.358*	–.81*
	(1.041)	(1.09)	(.395)
Initial information	–.11	–.458	.388
	(.532)	(.578)	(.442)
New information	.124	.201	.41
	(.439)	(.449)	(.366)
Involvement with the initial decision		.804	.85*
		(.517)	(.437)
Updated summary belief	.0936*	.092*	
	(.02)	(.0193)	

*$p < .05$
Notes: Standard errors are in parentheses.

---

[4] All 7 of the 98 participants in the high-involvement conditions who chose not to launch the product estimated probability of success to be less than 50%. The average for these seven participants was 46%, as opposed to 73% for those who continued. In addition, 5 of the 44 participants in low-involvement conditions forecasted a negative NPV and/or a less than 50% chance of a positive NPV. Given this initial negative forecast, we assumed that they would not have gone forward with the initial launch and thus eliminated them from further analysis. Furthermore, when we compared this initial decision not to launch for the two involvement groups, holding fixed the level of initial information, we observe that 5 of 59 did not launch in the high-involvement group, and we assume that 5 of 44 would not have launched in the low-involvement group. These proportions are not significantly different.

dicted value using the selected instruments, thus purging unobserved background factors from the measures. Note that the presence of correlation of our independent variables with unobserved individual differences is testable. Thus, in each equation, we used Hausman's (1978) test to identify whether the IV estimation procedure was required to avoid misleading estimates due to unobserved background factors that differed across participants and were correlated with both our dependent and independent measures. If the test indicated the presence of correlation, we used the IV procedure (this only happened for the summary-belief equation). As a result, our findings based on the use of measured variables can be interpreted as having the same validity as traditional analyses of variance based on manipulated variables.

Decision Involvement Distortion (H_2) suggests that involvement with the initial decision indirectly affects managers' decisions by altering the way they update their beliefs. If this is true, we would observe different patterns for the updating process for the two involvement groups. Consequently, we tested this hypothesis by determining whether the coefficient vectors, both slope and intercept, varied significantly between the two groups for each of the belief-structure updating equations.

We began by testing the summary-belief estimation equation (Equation 6). The appropriate F test indicated that the estimated intercept and slope coefficients did not vary between the two involvement groups for this equation ($F_{4,\,119} = .76$). Next, we estimated the updating equations for the three market-factor beliefs for the same two involvement groups (Equation 3). Again, we failed to find any evidence of a difference in updating between the two groups (market share: $F_{4,\,122} = 1.51$; competitive-entry response: $F_{4,\,127} = 1.19$; and market growth: $F_{4,\,122} = 1.53$). Finally, we estimated the updating equation for the importance weights associated with the market-factor beliefs. Here, we estimated a single updating equation (Equation 5) for the importance weights because the weights were constrained by a constant sum scale and therefore were interrelated. Again, we failed to find significant differences in updating for the two groups ($F_{3,\,362} = 1.14$). In summary, we failed to find any difference in updating of beliefs due to involvement with the initial decision, and therefore we found no support for the path associated with Decision Involvement Distortion.

Belief Inertia Distortion Results

Our previously reported logistic analysis suggests that updated summary beliefs are the driving forces behind the second continue/stop decision. Because we found that involvement does not affect the coefficients in the updating equations, we used the combined sample to estimate the three market-factor-belief-updating equations, the importance-weight-updating equation, and the summary-belief-updating equation. In each case, we tested H_3 by examining the nonnormative aspects of the equations.

Market-factor beliefs. We report the results for the market-factor-belief-updating equations in Table 2. These results indicate support for Belief Inertia Distortion. In support of H_3, we find that the effect of the initial summary belief is positive and significant in both the market-share-updating ($p < .05$) and the market-growth-updating ($p < .01$) equations. In addition, as we expected, the initial summary-belief effect is not significant in the competitive-entry-response-updating equation. Because there was no new information provided about the major competitor's entry response, participants would not be expected to engage in an updating process (biased or unbiased) regarding the competitive response.

Importance weights. We report the results from the importance-weight-updating equation in Table 3. Again, these results demonstrate support for Belief Inertia Distortion. Consistent with H_3, we found a significant ($p < .001$) increase (decrease) in importance weights for market factors when participants perceived the new information as

TABLE 2
Results of the Market-Factor-Belief-Updating Equations

	Market-Factor Belief		
Variable	**Market Share (N = 129)**	**Market Growth (N = 129)**	**Competitive-Entry Response (N = 128)**
Consistent with Normative Updating			
Intercept	3.047***	5.647***	3.032***
	(.592)	(.646)	(.464)
New information	.665**	–.088	–.169
	(.244)	(.258)	(.227)
Initial market-factor belief	.108	.149*	.348***
	(.075)	(.088)	(.078)
Consistent with Biased Updating			
Initial summary belief	.72*	.265**	–.084
	(.386)	(.102)	(.361)

*$p < .05$.
**$p < .01$.
***$p < .001$.
Notes Standard errors are in parentheses.

TABLE 3
Results of the Importance-Weights Equation
(N = 366)

Variable	
Consistent with Normative Updating	
Initial importance weight	.764* (.029)
Consistent with Biased Updating	
Relative positive or negative new component information	2.234* (.644)

*$p <$ 001.
Notes: Standard errors are in parentheses, no intercept regression is present in the model.

more (less) positive than their perception of the initial market-factor information.

Summary beliefs. We present the estimation results of the summary-belief-updating equation in Table 4. These results show that the effect of the initial summary belief is positive and significant ($p < .001$), as is the effect of integration of the participants' updated market-factor beliefs and importance weights ($p < .001$). As we noted in our discussion of Equation 6, for a well-calibrated manager, one but not both of these effects should be significant. Both terms are highly significant, which is a clear indication of bias entering the updated summary belief.

In summary, we find strong support for Belief Inertia Distortion. Moreover, the distortion appears concurrently at multiple levels—that is, at the aggregate level of summary beliefs and at the more disaggregated level of both component beliefs and importance weights.

Alternative Explanations for the Data

Previous theories of escalation, such as self-justification, face saving, and framing of options, have pointed to involvement with the initial decision as the main driver of escalation behavior. Our results are compatible with these previous theories in that we find that high-involvement decision makers are more likely to stay committed to a losing course of action than low-involvement decision makers. However, our results suggest that involvement with the initial decision is not a necessary condition for escalation behavior. Instead, we find that biased belief updating is the driving force behind escalation behavior. Indeed, we find that participants in the low-involvement conditions who formed initially favorable opinions of the new products exhibited the same behavior as participants in the high-involvement condition.

TABLE 4
Results of the Summary-Belief-Updating Equation
(N = 124)

Variable	
Intercept	34.664* (12.162)
New information	2.871 (30.151)
Initial summary belief^a	.251** (.066)
$\Sigma w_{ik2}FB_{ik2}{}^a$	.217** (.041)

*$p < .01$.
**$p < .001$.
^{a}Variable is instrumented.
Notes: Standard errors are in parentheses

Given the limited role of biased belief updating in previous studies, it might be expected that at least some, if not most, of the escalation effects would operate through the paths associated with decision involvement. However, our results counter much of the extant literature. Consequently, we explore possible alternative explanations for our findings.

Perhaps the simplest explanation for our finding of no effects from Decision Involvement Inertia and Decision Involvement Distortion is that our involvement manipulation did not provide meaningful separation between the two groups. This separation is critical because the explanations associated with these two paths rest on the assumption that the reevaluation decision makers differ in terms of involvement with the initial decision. Importantly, our three independent manipulation checks indicated significant differences between the two groups in terms of both the measured level of involvement and behaviors associated with different levels of involvement. Thus, we reject the idea that the lack of support for the explanations associated with these two involvement paths is due to a failed involvement manipulation.

Another possible alternative explanation for the lack of support for the decision involvement paths is the possibility of reverse causality in some of the paths. In this scenario, involvement/responsibility for the initial decision still acts as the "trigger" for escalation bias. The logic is as follows: Responsibility for the initial decision causes the decision maker to stay the course. Having made this decision, the decision maker modifies his or her beliefs to bring them into accord with the decision. That is, the causal arrow runs from the second decision to the updated beliefs.

To explore the feasibility of this account, we searched for evidence that would be compatible with this reverse causality argument. Specifically, the reverse causality interpretation implies that initial involvement with the decision must be correlated with the updated beliefs after we remove the mediating role of the subsequent decision. However, as we show in our analysis of the Decision Involvement Distortion path, the degree of involvement with the initial decision is independent of updated summary beliefs, market-factor beliefs, and weights. Therefore, the conjecture of reverse causality between the second decision and updated beliefs is incompatible with our data.

Implications

In this article, we provide an integrative framework that organizes the existing literature around three possible paths to escalation behavior. Although our methodology allows for the simultaneous conceptual and empirical existence of all three paths, our results suggest that the driving force behind escalation behavior in our study was not involvement with the initial decision but rather biased belief updating that outweighs initial positive beliefs.

In considering the theoretical implications of our findings, we acknowledge that these findings are based on one laboratory study, albeit with a diverse set of participants. Thus, although we believe that these findings generalize to other populations and other organizational settings (e.g., the National Basketball Association draft; hiring decisions; other strategic marketing decisions, such as selecting an ad campaign or choosing a new channel of distribution), we caution readers to draw their own conclusions about the extent to which our findings generalize. In addition, we acknowledge that all our participants had some involvement with the initial decision, and thus our study differed from some prior studies in which the noninvolved group had no contact whatsoever with the initial decision. That said, our findings suggest that Decision Involvement Inertia (involvement with the initial decision directly affecting subsequent decisions) is not the major driving force behind escalation. This casts doubt on several commonly invoked explanations for escalation, such as a framing effect, agency theory, and public or self-justification.

We acknowledge that this is a highly unexpected conclusion. Because the lack of a Decision Involvement Inertia effect was so surprising, we took extra care in our analysis to ensure that no such effect was being masked by either our method or our analysis. Having taken these steps, we point to two factors in our methodological approach that could account for why our findings differ from much of the existing escalation literature. First, with regard to escalation research, Northcraft and Wolfe (1984), Heath (1995), and Camerer and Weber (1999) point to the importance of having an explicit normative benchmark and note the absence of such benchmarks in most escalation studies. From this starting point, they conclude that though there are a large number of prior studies that examine escalation of commitment, these studies provide little empirical evidence in support of the existence of escalation bias. Therefore, we took great pains to provide information about marginal costs and benefits for each option, and in our setting, it is clear that the continuation decision is nonnormative. Thus, we believe that our study provides clear evidence of nonnormative escalation behavior. Second, Heath (1995) argues for the importance of disentangling the role and measurement of beliefs from the role and measurement of making the initial decision. Our study is the first to do this, enabling us to control for beliefs when testing for the effects of involvement with the initial decision.

We also find no support for the hypothesis of Decision Involvement Distortion (involvement with the initial decision leads to biased belief updating). Although this path has received much less attention in the literature, this result is important because it again shows that involvement with the initial decision is not the root cause for escalation behavior.

Instead, our results point to the powerful influence and complexity of biased belief updating, which weights initial positive beliefs too heavily (the Belief Inertia Distortion path). Note that the biasing influences on belief updating enter the process at many different levels. Specifically, we find that prior beliefs nonnormatively affect not only the updated market-factor beliefs but also the importance weights placed on those factors and the summary beliefs. Thus, all these sources of bias enter into the updating process concurrently.

Note also that these nonnormative updating results are compatible with Boulding, Kalra, and Staelin's (1999) results. This compatibility is notable because their setting was very different. In that study, consumers were the information processors, and longitudinal evaluations of service quality were used instead of updated evaluations of a new product introduction. Consequently, we conclude that the model of belief updating presented in our article should be useful in understanding belief updating in a wide range of situations. Our findings should be of interest to researchers who study belief formation and updating as well as those who are interested in understanding escalation. In general, the results point out that people have difficulty recognizing redundancies in the information they use to form overall beliefs. This is of particular importance to both marketing scientists who design marketing decision support systems and marketing researchers who are interested in understanding how consumers make significant decisions over time.

From a management perspective, the real issue is how to prevent escalation behavior in situations such as the introduction of a new product or the start of a new strategic alliance. Our findings should be helpful in determining which approaches have the most promise in preventing escalation behavior. However, because escalation was caused by a biased belief-updating process, our findings highlight the inherent difficulty in trying to eradicate escalation behavior completely. This may provide some explanation for why previous studies examining deescalation strategies produced mixed results (Boulding, Morgan, and Staelin 1997; Keil and Robey 1999; McNamara, Moon, and Bromiley 2002; Simonson and Staw 1992). The findings from those studies, in conjunction with our findings, lead us to suggest the following methods for preventing escalation behavior: (1) Change the organizational structure such that continue/stop decisions are made by someone with no prior beliefs about the project; (2) use stopping rules that are based on objective data; and (3) accept that decision makers hold biased beliefs, and institute policies and procedures that minimize the adverse effects of these biased beliefs. Such a policy could be an educational program that stresses the idea that managers tend to examine data with the goal of making the world appear consistent with their own views of reality (Carlson and Russo 2001). Specifically, it could stress the need for a decision maker to document how he or she might be wrong, thus forcing him or her to attend to negative information.

However, even these suggestions are not foolproof. For example, Bolton (2003) shows that such analytic reasoning is unlikely to eliminate bias from sticky priors if these priors were generated by nonanalytic thinking. Likewise, when McNamara, Moon, and Bromiley (2002) tried to use the approach of changing decision makers, they were able to attenuate but not eliminate escalation in a commercial lending situation. Similarly, Boulding, Morgan, and Staelin (1997) found that stopping rules adopted at the time of the initial decision also attenuated escalation but that some decision makers simply ignored their own stated stopping rule when making subsequent decisions. This willingness to change the rules is similar to what Keil, Mann, and Rai (2000) found in the domain of information technology projects. They found that the best explanation for escalation in such projects was what they referred to as the "completion effect"—a tendency to substitute the goal of completing the project for the goal of maximizing economic returns. Together, these studies indicate that implementing organizational solutions will not be easy and that managers need to be doubly alert to the possibility of escalation bias. Moreover, research is needed to understand the underlying drivers of biased belief updating to help identify better ways to reduce escalation behavior in managerial settings. For example, research on decision processes at the team or organization level may point to additional factors that might help individuals and, therefore, firms avoid the trap of escalation bias.

Appendix

Manipulation Check 1[5]

In the first manipulation check, we collected data from 59 participants from the same MBA student pool previously described, varying only the high-involvement (29 participants) and low-involvement (30 participants) conditions. Participants completed the same initial decision/evaluation part as in the main study. After reading the part in the case that informed them that the company decided to proceed with the product introduction, they were asked to indicate how well ten different questions (see the "Manipulation Check Questions") reflected their beliefs. The questions assessed (on a nine-point scale ranging from "low" to "high") how much respondents felt publicly committed, accountable, and responsible for the product launch.[6]

Each of the ten items is significantly different ($p < .001$ or better) between the low- and the high-involvement groups and in the expected direction. If we sum the items and divide by ten, the mean of the low-involvement group is 5.1 and the mean of the high-involvement group is 7.2. These means are significantly different ($t = 6.84$, $p < .001$), and these results indicate significantly different involvement in the low- and high-involvement groups.

Manipulation Check 2

A total of 44 participants taken from the same MBA student pool and evenly divided between the high- and the low-involvement conditions completed the same initial decision/evaluation part as in the main study. After finishing this decision/evaluation task, they were asked to fill in a debriefing sheet that instructed them to "describe why you think the Quality Valve Company should or should not launch the new product."

Prior research indicates that people who feel accountable for (i.e., psychologically and/or socially involved with) a decision have a need to justify their position. Therefore, they generate attitude-consistent thoughts, which lead to more (less) thoughts that support (oppose) their initial attitude (Huber and Seiser 2001; Lambert et al. 1996; Schlenker 1980; Tetlock 1985; Tetlock, Skitka, and Boettger 1989). Thus, we expect the high-involvement respondents to mention more arguments in support of and fewer arguments opposed to their decision than the low-involvement respondents. In addition, evidence in the literature suggests that people who are accountable for their decisions write more thorough justifications (Koonce, Anderson, and Marchant 1995), are more likely to qualify their opinions, and exhibit more recognition of the trade-offs involved (Lord 1992). Thus, we expect the respondents in the high-involvement condition to report relatively more arguments that make explicit trade-offs between the pluses and minuses of the decision.

Two coders, who were blind to the hypotheses, placed participants' arguments into the following three categories: (1) arguments that support the initial decision,[7] (2) arguments that do not support the initial decision, and (3) arguments that make explicit trade-offs between supporting and nonsupporting aspects. The percentage of agreement between the coders was 84%, with a Cohen's (1960) kappa of .73 and a Perreault and Leigh (1989) reliability index of .87 (95% confidence interval of .81–.93). All measures indicate a reasonable level of agreement between the two coders.

We used the F test for differences in count data (Kanji 1993, p. 51) to test for differences between our manipulated conditions. As we expected, the number of supporting arguments made by the high-involvement group was significantly greater than the number made by the low-involvement group ($F_{67, 110} = 1.62$, $p < .05$), and the pattern was reversed for nonsupporting arguments ($F_{29, 54} = 1.77$, $p < .05$).[8] As we also expected, the high-involvement group produced significantly more explicit trade-off arguments than the low-involvement group ($F_{3, 33} = 8.00$, $p < .001$).

Manipulation Check 1 Questions

Respondents were asked to assess "How well do the following statements reflect your beliefs" on a nine-point scale ranging from "completely disagree" to "completely agree" for the following ten statements:

1. If the project succeeds, I will be given substantial credit.
2. People will associate the success or failure of the project with me.

---

[5] In both manipulation check studies, we collected data in the positive prior information/negative new information conditions.

[6] Cronbach's alpha for the ten items is .94. Principal component factor analysis indicates a single factor solution.

[7] The coders actually coded arguments as for or against the launch. For obvious reasons, we did not want the coders to be aware of the respondent's actual decision. By examining participants' launch decisions later, we were able to ascertain whether the arguments were supporting or not.

[8] All reported results are based on the average of the two coders' values. Analysis based on each coder separately yields the same results.

3. I am largely responsible for the decision to launch the product.
4. The outcome of this project will have a major impact on my career.
5. If the project fails, I will be given substantial blame.
6. People will attribute the decision to me.
7. I feel responsible for the chosen course of action.
8. I feel attached to the decision to launch the product.
9. My input into the decision process was public.
10 If the project fails, top management will believe that I gave poor advice.

REFERENCES

Ajzen, I. and M. Fishbein (1977), "Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research," *Psychological Bulletin*, 84 (5), 233–50.

Arkes, H.R. and C. Blumer (1985), "The Psychology of Sunk Costs," *Organizational Behavior and Human Decision Processes*, 35 (1), 124–40.

Bacharach, S.B., P. Bamberger, and B. Mundell (1995), "Strategic and Tactical Logics of Decision Justification. Power and Decision Criteria in Organizations," *Human Relations*, 48 (5), 467–88

Barton, S.L., D. Duchon, and K.J. Dunegan (1989), "An Empirical Test of Staw and Ross's Prescriptions for the Management of Escalation of Commitment Behavior in Organizations," *Decision Sciences*, 20 (3), 532–44.

Bazerman, M. (1984), "The Relevance of Kahneman and Tversky's Concept of Framing to Organizational Behavior," *Journal of Management*, 10 (3), 333–43.

——— and M. Neale (1992), "Nonrational Escalation of Commitment in Negotiation," *European Management Journal*, 10 (2), 163–68.

Bettman, J.R. (1979), *An Information Processing Theory of Consumer Choice*. Reading, MA: Addison-Wesley.

Bobocel, D.R. and J.P. Meyer (1994), "Escalating Commitment to a Failing Course of Action· Separating the Roles of Choice and Justification," *Journal of Applied Psychology*, 79 (3), 360–33.

Bolton, Lisa E. (2003), "Stickier Priors: The Effects of Nonanalytic Versus Analytic Thinking in New Product Forecasting," *Journal of Marketing Research*, 40 (February), 65–79

Boulding, W., A. Kalra, and R. Staelin (1999), "The Quality Double Whammy," *Marketing Science*, 18 (4), 463–84.

———, Marian Chapman Moore, Richard Staelin, Kim P. Corfman, Peter Reid Dickson, Gavan Fitzsimons, Sunil Gupta, Donald R. Lehmann, Deborah J. Mitchell, Joel E. Urbany, and Barton A. Weitz (1994), "Understanding Managers' Strategic Decision-Making Process," *Marketing Letters*, 5 (4), 413–26.

———, Ruskin Morgan, and Richard Staelin (1997), "Pulling the Plug to Stop the New Product Drain," *Journal of Marketing Research*, 34 (February), 164–76.

Brockner, J. and J.Z. Rubin (1985), *Entrapment in Escalating Conflicts*. New York: Springer-Verlag.

———, ———, and E. Lang (1981), "Face-Saving and Entrapment," *Journal of Experimental Social Psychology*, 17 (1), 68–79.

Camerer, C.F. and R.A. Weber (1999), "The Econometrics and Behavioral Economics of Escalation of Commitment: A Re-Examination of Staw and Hoang's NBA Data," *Journal of Economic Behavior and Organization*, 39 (1), 59–82.

Carlson, K.A. and J.E. Russo (2001), "Biased Interpretation of Evidence by Mock Jurors," *Journal of Experimental Psychology-Applied*, 7 (2), 91–103.

Cohen, J. (1960), "A Coefficient of Agreement for Nominal Scales," *Educational and Psychological Measurement*, 20 (1), 37–46.

Conlon, D.E. and H. Garland (1993), "The Role of Project Completion Information in Resource Allocation Decisions," *Academy of Management Journal*, 36 (2), 402–413.

Einhorn, H.J. (1970), "The Use of Nonlinear, Noncompensatory Models in Decision Making," *Psychological Bulletin*, 73 (2), 221–30.

Festinger, L. (1957), *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.

Fox, F.V. and B.M. Staw (1979), "The Trapped Administrator: Effects of Job Insecurity and Policy Resistance upon Commitment to a Course of Action," *Administrative Science Quarterly*, 24 (3), 449–71.

Hausman, J.A. (1978), "Specification Tests in Econometrics," *Econometrica*, 46 (6), 1251–71.

Heath, C. (1995), "Escalation and De-Escalation of Commitment in Response to Sunk Costs: The Role of Budgeting in Mental Accounting," *Organizational Behavior and Human Decision Processes*, 62 (1), 38–54.

Hoch, S.J. and Y. Ha (1986), "Consumer Learning: Advertising and the Ambiguity of Product Experience," *Journal of Consumer Research*, 13 (2), 221–33.

Hsee, C.K. (1995), "Elastic Justification· How Tempting but Task-Irrelevant Factors Influence Decisions," *Organizational Behavior and Human Decision Processes*, 62 (3), 330–37.

Huber, O. and G. Seiser (2001), "Accounting and Convincing: The Effect of Two Types of Justification on the Decision Process," *Journal of Behavioral Decision Making*, 14 (1), 69–85.

Johnson, E. and R. Meyer (1984), "Compensatory Models of Non-Compensatory Choice Processes: The Effect of Varying Context," *Journal of Consumer Research*, 11 (1), 528–41.

Jonas, E., S. Schulz-Hardt, D. Frey, and N. Thelen (2001), "Confirmation Bias in Sequential Information Search After Preliminary Decisions: An Expansion of Dissonance Theoretical Research on Selective Exposure to Information," *Journal of Personality and Social Psychology*, 80 (4), 557–71.

Kahneman, D. and A. Tversky (1979), "Prospect Theory: An Analysis of Decisions Under Risk," *Econometrica*, 47 (2), 263–91.

Kanji, G.K. (1993), *100 Statistical Tests*. London: Sage Publications.

Keeney, R.L. and H. Raiffa (1976), *Decision with Multiple Objectives: Preferences and Value Tradeoffs*. New York: John Wiley & Sons.

Keil, M. (1995), "Escalation of Commitment in Information Systems Development: A Comparison of Three Theories," *Academy of Management Journal*, 38 (2), 348–52.

———, J. Mann, and A. Rai (2000), "Why Software Projects Escalate: An Empirical Analysis and Test of Four Theoretical Models," *MIS Quarterly*, 24 (4), 631–64.

——— and D. Robey (1999), "Turning Around Troubled Software Projects: An Exploratory Study of the Deescalation of Commitment to Failing Courses of Action," *Journal of Management Information Systems*, 15 (1), 63–87.

Klayman, J. (1995), "Varieties of Confirmation Bias," in *Decision Making from a Cognitive Perspective: The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 32, J. Busemeyer, R. Hastie, and D.C. Medin, eds. San Diego: Academic Press.

Koonce, L., U. Anderson, and G. Marchant (1995), "Justification of Decisions in Auditing," *Journal of Accounting Research*, 33 (Autumn), 369–84.

Lambert, A.J., S. Cronen, A.L. Chasteen, and B. Lickel (1996), "Private Versus Public Expressions of Racial Prejudice," *Journal of Experimental Social Psychology*, 32 (5), 437–59.

Lipshitz, R. (1991), "The Road to Desert Storm," *Organization Studies*, 16 (2), 242–64.

Lord, A.T. (1992), "Pressure: A Methodological Consideration for Behavioral Research in Auditing," *Auditing: A Journal of Practice and Theory*, 11 (2), 90–108.

Lukas, P. (1998), "The Ghastliest Product Launches," *Fortune*, (March 16), 44.

McAllister, D.W., T.R. Mitchell, and L.R. Beach (1979), "The Contingency Model for the Selection of Decision Strategies: An Empirical Test of the Effects of Significance, Accountability, and Reversibility," *Organizational Behavior and Human Performance*, 24 (2), 228–244.

McNamara, G., H. Moon, and P. Bromiley (2002), "Banking on Commitment: Intended and Unintended Consequences of an Organization's Attempt to Attenuate Escalation of Commitment," *Academy of Management Journal*, 45 (2), 443–52.

Mitroff, I. (1974), *The Subjective Side of Science*. Amsterdam: Elsevier.

Northcraft, G.B. and G. Wolf (1984), "Dollars, Sense, and Sunk Costs: A Life Cycle Model of Resource Allocation Models," *Academy of Management Review*, 9 (2), 225–34.

Perreault, William D., Jr., and Laurence E. Leigh (1989), "Reliability of Nominal Data Based on Qualitative Judgments," *Journal of Marketing Research*, 26 (May), 135–48.

Ross, J. and B.M. Staw (1986), "Expo 86: An Escalation Prototype," *Administrative Science Quarterly*, 31 (2), 274–97.

Samuelson, W. and R. Zeckhauser (1988), "Status Quo Bias in Decision Making," *Journal of Risk and Uncertainty*, 1 (1), 7–59.

Schlenker, B.R. (1980), *Impression Management: The Self-Concept, Social Identity, and Interpersonal Relations*. Monterey, CA: Brooks/Cole Publishing.

Schmidt, J.B. and R.J. Calantone (1998), "Are Really New Product Development Projects Harder to Shut Down?" *Journal of Product Innovation Management*, 15 (2), 111–23.

Schwenk, C.R. (1986), "Information, Cognitive Biases, and Commitment to a Course of Action," *Academy of Management Review*, 11 (2), 298–310.

Silver, W.S. and T.R. Mitchell (1990), "The Status Quo Tendency in Decision Making," *Organizational Dynamics*, 18 (4), 34–46.

Simonson, I. and B.M. Staw (1992), "Deescalation Strategies: A Comparison of Techniques for Reducing Commitment to Losing Courses of Action," *Journal of Applied Psychology*, 77 (4), 419–26.

Staw, B.M (1976), "Knee-Deep in the Big Muddy: A Study of Escalating Commitment to a Chosen Course of Action," *Organizational Behavior and Human Performance*, 16 (1), 27–44.

——— and H. Hoang (1995), "Sunk Costs in the NBA: Why Draft Order Affects Playing Time and Survival in Professional Basketball," *Administrative Science Quarterly*, 40 (3), 474–94.

——— and J. Ross (1987), "Behavior in Escalation Situations: Antecedents, Prototypes, and Solutions," in *Research in Organizational Behavior*, Vol. 7, B.M. Staw and L.L Cummings, eds Greenwich, CT: JAI Press.

Tetlock, P.E (1985), "Accountability: The Neglected Social Context of Judgment and Choice," *Research in Organizational Behavior*, 7, 297–332.

———, L. Skitka, and R. Boettger (1989), "Social and Cognitive Strategies for Coping with Accountability: Conformity, Complexity, and Bolstering," *Journal of Personality and Social Psychology*, 57 (4), 632–40.

Whyte, G. (1986), "Escalating Commitment to a Course of Action: A Reinterpretation," *Academy of Management Review*, 11 (2), 311–21.

Wicklund, R.A. and J.W. Brehm (1976), *Perspectives on Cognitive Dissonance*. Hillsdale, NJ. Lawrence Erlbaum Associates.

Michael D. Johnson, Andreas Herrmann, & Frank Huber

The Evolution of Loyalty Intentions

The drivers of customer loyalty intentions are dynamic. What remains unclear is how these intentions evolve through the introduction and growth phases of a life cycle. Using a longitudinal study of cellular phone customers, the authors demonstrate that loyalty intentions are a function of perceived value early in the life cycle. Over time, more affective attitudes toward the brand and the relationship with the company come to mediate the effects of value on intentions. The results suggest that from the introduction to the growth stage of a life cycle, managers must adapt from improving value per se to measuring and managing relationships and brands directly.

Providing customers with perceived value or customer satisfaction is widely recognized as a means of improving loyalty intentions (Fornell et al. 1996; Zeithaml, Berry, and Parasuraman 1996) and actual retention (Bolton 1998; Bolton and Lemon 1999; Mittal and Kamakura 2001). However, research demonstrates that these relationships are potentially complex and dynamic and that the drivers of intentions change and evolve over time (Mittal, Kumar, and Tsiros 1999; Slotegraaf and Inman 2004). Such research focuses on relatively mature product or service categories, such as automobiles and credit cards. What remains unclear is how the drivers of intentions evolve for a new-to-market offering.

A better understanding of these dynamics is essential for researchers and managers alike. It enables researchers to focus on the most important predictors of performance while helping managers design marketing programs that effectively adapt to customers through the diffusion process. This article builds on dynamic marketing models in two important ways: First, we explore the drivers of intentions for a panel of loyal customers who purchased and used a new-to-market offering—specifically, a new-generation cellular phone technology—through the introduction and growth phases of a life cycle. Second, we incorporate two particular attitudes that have not been studied previously in this line of research: the affective commitment toward maintaining a relationship with the company (Fullerton 2003; Garbarino and Johnson 1999; Verhoef 2003) and brand equity (Keller 1993; Rust, Zeithaml, and Lemon 2000).

We first describe research on the dynamics of loyalty intentions and then develop our conceptual model and hypotheses. Using attitude theory, we argue that perceived value, as an overall evaluation of performance given price or prices paid, has a more direct influence on customers' intentions early in a life cycle when customers are gathering information, forming opinions about brands and relationships, and trying to reduce risk (Heilman, Bowman, and Wright 2000). Over time, attitudes toward the brand and relationship should come to mediate the effect of value on intentions (Oliver 1999). We test our model on a panel of cellular phone customers who were loyal to a particular brand over a four-year period that coincided with the introduction and growth phases of the life cycle. The results illustrate how loyalty intentions evolve from being purely value based early in the diffusion process to being brand and relationship based as the market evolves. The study has important implications for managers who are charged with allocating resources to improve product benefits, brand perceptions, and customer relationships over time.

The Evolution of Loyalty Intentions

The vast majority of work relating perceived performance to intentions has been cross-sectional. However, recent research has suggested that intention drivers vary considerably over time. Mittal, Kumar, and Tsiros (1999) compare the drivers of perceived performance and subsequent loyalty intentions for new vehicle owners 3 to 4 months after purchase and again 21 months later. They find that service satisfaction has a greater impact on intentions early in the relationship, whereas product satisfaction has a greater impact later in the relationship. In a subsequent study, Mittal, Katrichis, and Kumar (2001) examine the drivers of customer satisfaction for people who owned a credit card for less than a year versus more than a year. They find that the weight customers place on different performance attributes shifts on the basis of their tenure with a firm. Again, in an automotive context, Slotegraaf and Inman (2004) study changes in the effects of resolvable attributes (attributes that can be fixed or repaired under warranty, such as a faulty part) versus irresolvable attributes (attributes that cannot be fixed or changed regardless of the warranty, such as size or gas mileage). They find that resolvable attributes have greater influence on satisfaction over time, whereas irresolvable attributes have less influence.

These studies reinforce the need for a dynamic perspective. Yet the settings for these studies involve relatively mature product categories (e.g., automobiles, credit cards). Logically, loyalty intentions should be particularly dynamic for a new-to-market offering. Theory suggests a specific dynamic for such offerings, for which more affective-based

---

Michael D. Johnson is D. Maynard Phelps Professor of Business Administration, Stephen M. Ross School of Business, University of Michigan (e-mail: mdjohn@umich.edu). Andreas Herrmann is Professor of Business Metrics, ZBM Institute, University of St. Gallen (e-mail: andreas.herrmann@unisg.ch). Frank Huber is Professor of Marketing, University of Mainz (e-mail: huber@marketing-mainz.de). The authors thank Michel Wedel for his comments on a previous draft of the article and Frank Kressmann for his assistance with the analysis.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 70 (April 2006), 122–132

attitudes come to mediate the effects of performance perceptions on intentions over time. Following traditional attitude models (Bagozzi and Warshaw 1990), we expect that the degree of mediation from perceived value, to brand and relationship attitudes, to intentions varies systematically as the market evolves. As the customers' experience with a particular offering grows over time, attitudes toward brands and relationships should become stronger, more "top of mind" or accessible, persistent, resistant to change, and likely to guide intentions and subsequent behavior (Fazio, Powell, and Williams 1989; Priester et al. 2004).

We use customers' overall perceived value to capture perceptions of performance. As we noted previously, two important attitude-type constructs have emerged in the literature as possible mediators of the value–intentions relationship. The first is affective commitment (Morgan and Hunt 1994; Verhoef 2003), or the factors that create stickiness in a relationship. This commitment is affected directly by the degree of personal interaction between a customer and a company and how the company manages the customer's account over time (Bendapudi and Leone 2002). The second is brand equity, which is influenced by factors such as repeated performance and satisfaction (Keller 1993; Selnes 1993), word of mouth or the "buzz" about the brand (Rosen 2000), the degree of identification with the brand, and its relevance to a customer's situation (Aaker 2004; Keller 2003).

Perceived Value

Perceived value is a customer's overall evaluation of what he or she receives compared with what he or she gives up or pays (Bolton and Drew 1991). Similar constructs in the marketing literature include payment equity (Bolton and Lemon 1999), distributive equity or fairness (Olsen and Johnson 2003), and reciprocity (Bagozzi 1975). Following the work of Bolton and Lemon (1999), we conceptualize perceived value as a broad construct that encompasses perceptions of quality given price and inputs versus outputs relative to the competition. We do not include overall customer satisfaction in our study. Empirically, however, perceived value and customer satisfaction are closely related constructs in the literature (Bolton and Lemon 1999; Fornell et al. 1996). Moreover, when customers are in an information-gathering or problem-solving mode, as should be the case in the early stages of a life cycle, value or equity perceptions are more relevant (Olsen and Johnson 2003).

Affective Commitment

Research on relationship commitment distinguishes between affective commitment and calculative commitment (Bendapudi and Berry 1997; Hansen, Sandvik, and Selnes 2003; Johnson et al. 2001). Our focus is on affective commitment, which is a "hotter" or more emotional factor related to the degree to which a customer identifies and is personally involved with a company and the resulting degree of trust and commitment (Bendapudi and Berry 1997; Garbarino and Johnson 1999; Morgan and Hunt 1994). Similar constructs used to describe exchange relationships include friendship, rapport, and trust (Fullerton 2003). Initially developed to explain employee commitment to organizations or work groups, the concept applies to consumption relationships as well (Gruen, Summers, and Acito 2000). In a recent study of financial services, Verhoef (2003) demonstrates direct effects of affective commitment on actual behavior, both relationship maintenance (retention) and relationship development (share of a customer's business). Although both perceived value or payment equity and customer satisfaction were positive antecedents of affective commitment, they did not directly affect behavior.

Calculative commitment, also called continuance commitment (Fullerton 2003), captures the "colder" or more rational, economic-based dependence on product benefits that is due to a lack of choice or switching costs that make it difficult to change suppliers (Anderson and Weitz 1992; Dwyer, Schurr, and Oh 1987). The company survey we use in this study contains only measures of affective commitment. However, we expect that some significant calculative commitment or switching costs exist and influence behavior, including the cost of the cellular phone itself and the inability of customers to transfer their numbers from phone to phone (Gustafsson, Johnson, and Roos 2005).

Brand Equity

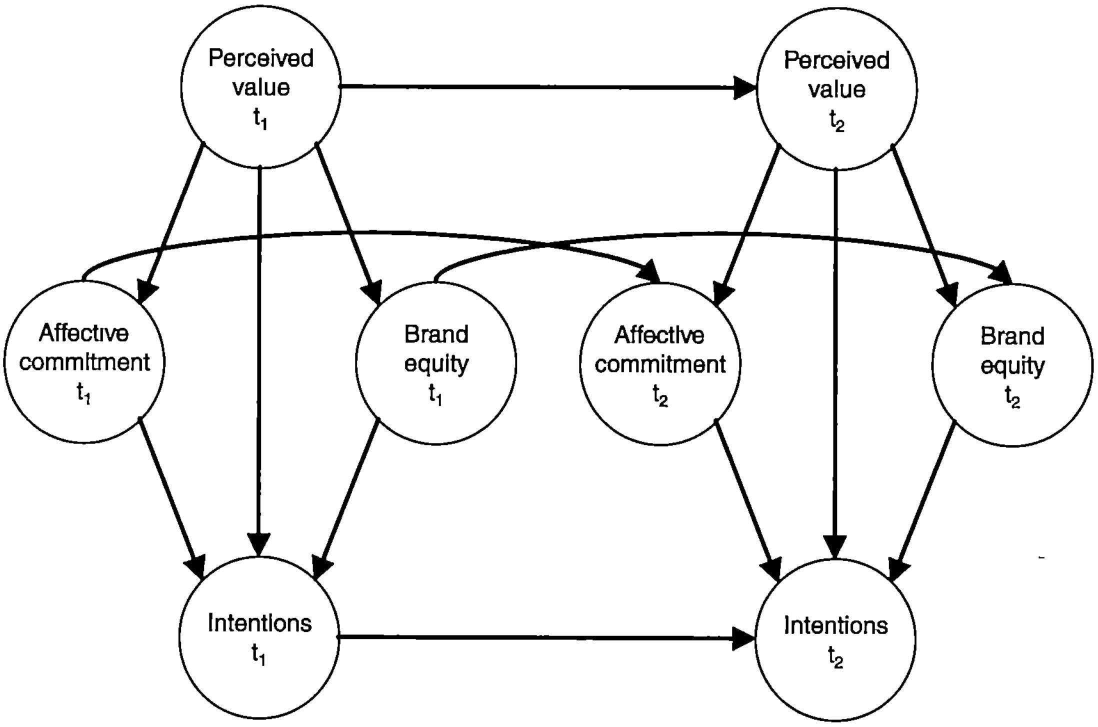
Brand equity is defined as "the differential effect of brand knowledge on consumer response to the marketing of a brand" (Keller 1993, p. 8). Central to this definition is the concept of differential effect. When loyalty intentions are explained, this is the effect that brands have beyond the rest of the value proposition. In our conceptual model, which we describe subsequently, the two major non-brand-related effects are overall perceived value (products and services received for price or prices paid) and affective commitment, or the strength of the relationship with a company. Research shows that brand equity has a direct effect on intentions and mediates the effects of quality and satisfaction on intentions to various degrees (Selnes 1993).

Particularly relevant to brand equity is the degree of personal identification with a brand and its relevance to a customer's situation (Aaker 2004; Keller 2003). It is this identification or personal "fit" with a brand that helps create the differential effect of brand equity on intentions beyond the effects of performance or relationships per se. Keller (1993) describes two approaches to measuring brand equity: indirectly through brand knowledge (brand awareness and image) and directly through the impact of brand knowledge on customer responses. Our study uses elements of both approaches. We measure brand equity as a latent variable using multiple measures, such as the degree to which the brand fits a customer's personality and lifestyle. We then determine whether the construct explains loyalty intentions beyond the influences of perceived value and affective commitment.

Conceptual Model

Our conceptual model in Figure 1 posits that perceived value builds loyalty intentions both directly and indirectly through the creation of brand equity and affective commitment. The prediction that brand equity has greater influence on intentions over time is consistent with other research on the evolution of brand preferences. Heilman, Bowman, and

FIGURE 1
The Evolution of Loyalty Intentions: Conceptual Model



Wright (2000) examine how brand preferences and responses to marketing activities evolve for customers who are new to established product categories (disposable diapers and towels). They find little evidence of brand loyalty early in customers' experiences, but as customers gather more information and as perceived risk subsides, they buy the brands they prefer.

Our conceptual model is consistent with Rust, Zeithaml, and Lemon's (2000) customer equity model, which includes three categorically different ways to increase customer equity: value equity, brand equity, and retention equity. Value equity is a customer's relatively cognitive perception of value, brand equity includes subjective appraisals of a brand that are not captured by objective performance, and retention equity is that which is gained from relationship-building and retention programs. For example, using data from the airline industry, Rust, Lemon, and Zeithaml (2004) show the projected returns to an organization from investments in various customer equity drivers (e.g., quality, advertising, loyalty programs). Our model is also consistent with the logic that underlies Berry and Parasuraman's (1991) framework for customer retention strategies. Accordingly, early in a relationship, customers are tied to a firm primarily through product or service benefits and financial incentives (perceived value). Stronger relationships grow through personal relationships, social bonds, and customer intimacy.

Temporal Effects and Consumption Systems

Prior studies have demonstrated that the drivers of intentions are dynamic; that is, they vary over time. What these studies do not show is how these drivers evolve as customers repeatedly use and experience an offering through the introduction and growth stages of a life cycle. An important factor to consider when modeling this evolution is the existence of temporal or carryover effects from one period to another. Early on, Oliver (1980) demonstrated the time dependence of attitudes and behavioral intentions in a satisfaction model. He showed how attitudes in one period are a function of both customer satisfaction in that period and attitudes in the previous period. In turn, behavioral intentions in a given period are a function of attitudes in that period and prior period intentions. Similarly, Bolton and Drew (1991) find temporal effects for both attitudes and overall quality perceptions.

The time dependence of evaluations, attitudes, and intentions is a central feature of Mittal, Kumar, and Tsiros's (1999) consumption-system approach. They define (p. 89) a consumption system as "a bundle of goods and services that are consumed over time in multiple consumption episodes." Temporal effects in a consumption system are similar to the effects in Oliver's (1980) and Bolton and Drew's (1991) studies, in which perceptions, attitudes, and intentions in one period serve as anchors for the same constructs in subsequent periods. Evaluations of perceived value, brand equity, affective commitment, and loyalty intentions are not constructed anew each period. Rather, they are updated versions of prior evaluations (Helson 1948; Oliver 1980).

Another feature of the consumption-system approach is crossover effects, or the degree to which perceived performance in one part of the system (as with the product) affects behavioral intentions in another part of the system (as with the service) over time. Because our data come from the cellular phone provider and focus on loyalty intentions to the phone, we focus on capturing temporal effects rather than crossover effects in the consumption-system approach. Our conceptual model incorporates the temporal effects of evaluations of value, brand equity, affective commitment, and loyalty intentions in one period on the same evaluations in a subsequent period. Because the model controls for temporal effects in the system, we gain a better understanding of the influence of perceived value, brands, and relationships on intentions.

Research Hypotheses

To summarize, we predict that perceived value has a more direct effect on intentions early in a life cycle. As the life cycle evolves, more affective attitudes, captured by the brand equity and affective commitment constructs in our model, should have greater influence. Moreover, the brand and relationship constructs should come to mediate the effects of perceived value on intentions. By studying an evolving market, we are in a unique position to gain an understanding of how loyalty drivers grow and change as competition grows.

We state these predictions formally as hypotheses. Our first hypothesis is related to the direct effect of perceived value on intentions over time. We expect that perceived performance has a direct effect on intentions that decreases over time as other factors (i.e., brands and relationships) come to influence intentions.

> H_1: The perceived value of an offering has a direct, positive effect on loyalty intentions that decreases as a market grows.

As part of the mediation argument, we further predict that an offering's perceived value should build brand equity and affective commitment.

> H_2: The perceived value of an offering has a direct, positive effect on brand equity that increases as a market grows.
>
> H_3: The perceived value of an offering has a direct, positive effect on affective commitment that increases as a market grows.

As we argue in H_1, the direct effect of perceived value on loyalty intentions should decrease over time as the brand and relationship constructs grow and come to influence intentions. H_4 and H_5 capture the increase in the effects of brand equity and relationship commitment on loyalty intentions that are consistent with the mediation argument:

> H_4: Brand equity has a direct, positive effect on loyalty intentions that increases as a market grows.
>
> H_5: Affective commitment has a direct, positive effect on loyalty intentions that increases as a market grows.

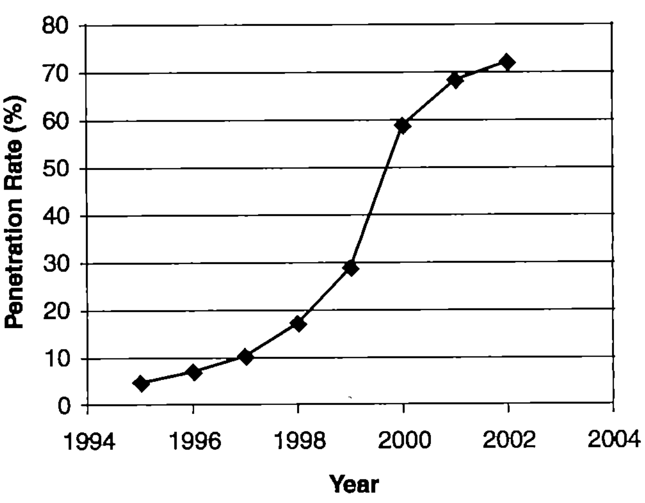
Although not stated as formal hypotheses, consistent with Mittal, Kumar, and Tsiros's (1999) findings, we expect significant temporal effects in the consumption system. That is, perceived value, brand equity, affective commitment, and loyalty intentions in one period should affect the same construct in a subsequent period. All our predictions are time dependent. As a market continues to mature, the influence of perceived value in relation to brands and relationships should stabilize. As we describe in the next section, because we estimate the model over three periods, there are a relatively large number of measures, constructs, and relationships involved. Given the possible number of alternative specifications, we do not posit and test any alternative models per se. Rather, we implicitly examine alternative specifications through our analysis of the size and significance of the path coefficients and how they change over time. We also conduct formal tests of mediation among our constructs (Baron and Kenny 1986).

Empirical Study of the Cellular Phone Market

We test the hypotheses using data from a commercial tracking study conducted for a global supplier of mobile communications in the German market. The study was fielded at three points in time that were approximately two years apart: September 1996, August 1998, and October 2000. Figure 2 shows the penetration rate over time for the cellular phone market that was surveyed. The 1996 and 1998 surveys coincide with the introductory and early growth phases of the market. By October 2000, the market had witnessed dramatic growth and an increase in competition.

First-generation cellular phones were simply devices for making calls on a single network. The second-generation cellular phones, which we study here, represent a new-to-market offering in several respects, including the use of multiband technology (users are able to lock into multiple networks), short message system capability, a Web interface, a significantly lighter weight phone, longer battery life, and various service features. Several fundamental changes in the German cellular phone market stimulated the development and diffusion of this new offering. Before 1989, Deutsche Telekom (DT) monopolized the market, which limited market penetration. From 1989 through 1998, there was a stepwise withdrawal of government influence on DT and a reduction in entry barriers for competitors. On August 1, 1996, the market was officially opened for any network provider. By 1998, four major network providers

FIGURE 2
Penetration Rate by Year for Cellular Phones in Germany



(DT, Mannesmann, E-Plus, and Viag Intercom) were collaborating with the cellular phone hardware providers (Nokia, Samsung, Motorola, Siemens, and Sony-Ericsson) to provide product/service offerings. Although the network and hardware providers were interchangeable, our data focus on the hardware provider.

The original 1996 sample included 5486 respondents. The sample size was reduced because of panel mortality to 4234 respondents in 1998 and to 3148 respondents in 2000. Another 248 respondents were removed from the sample because of high levels of missing data and outliers. This left a sample of 2990 respondents with usable responses in each period; each respondent owned a cellular phone and was interviewed in person. Of these customers, 1472 (49%) were loyal to the same brand of phone in all three periods. It is important for our analysis to focus on these brand-loyal customers for two reasons: First, theoretically, our goal is to capture the evolution of loyalty intentions for just such a group. Second, empirically, the carryover effects in the consumption system are more meaningful when the same brand is involved. The market shares of the different brands owned by the sample in each wave are similar to the shares of each brand in the marketplace. Of the 1472 loyal customers, 63% were male (37% were female), 60% had at least one child at home, and 31% had some university education. Of 2990 loyal and nonloyal customers, 68% were male, 48% had a child at home, and 30% had some university education. This indicates that the loyal customer sample includes somewhat more females and at least one child at home. Income and age information is available only by ranges in the survey, but in each case, the distributions were almost identical.

By definition, behavioral loyalty or retention is invariant across our test sample, making intentions the ultimate dependent variable in our model. However, these intentions may not be related to actual behavior (Verhoef 2003). A simple analysis examines whether our intentions are related to behavior. Beginning with our entire sample of 2990 usable responses in Year 1 (1996), we extracted the first principal component from the loyalty intention measures. An analysis of variance reveals that the intentions of the loyal customers in our test sample were indeed higher in the first wave of the survey than were the intentions of those who subsequently switched (principal component values of .043 versus –.041; $F = 5.285$, $p < .05$).

There are both advantages and disadvantages associated with using the same sample over time. The important theoretical advantage is that time translates directly into accumulated customer experience, resulting in risk reduction for the brand-loyal customers in our test sample. The disadvantage is that we focus on what may be a very particular customer population, namely, technology upgraders or early adopters who readily embrace new technologies (Parasuraman 2000). Differences observed over time may be due to the loyal sample simply being different from other customers. Arguably, the advantages of using the sample outweigh the disadvantages. First, our theoretical arguments rest on the assumption that customers remain loyal to a particular cellular phone. Second, there is no readily available argument to suggest why the intention drivers would evolve as predicted simply because this particular sample is involved. However, given the potentially unique nature of the sample, it is important that we focus on relative changes in intention drivers over time.

Survey Measures

The survey measures we used to estimate the conceptual model appear in Table 1. Respondents rated all measures on scales ranging from 1 ("completely disagree") to 7 ("completely agree"). The perceived value measures include the performance received for the price paid, whether the offering was a good deal compared with others in the marketplace, whether the price was fair, and whether the offering was a great value overall. We measured brand equity, affective commitment, and loyalty intentions using five survey measures each. The brand equity measures include whether the brand reflects customers' personal lifestyles, whether the brand fits their personality, and brand identification. The affective commitment measures include customers' commitment to maintaining a relationship with the manufacturer and personal interactions with the manufacturer (provision of feedback, interactive events, and other dialogue with the manufacturer). Finally, the loyalty intentions include repurchase intentions and recommending the cellular phone to others.

Model Specification, Estimation, and Evaluation

We specify all the constructs in the model using reflective indicators from Table 1 (Fornell and Cha 1994). The estimation method should be consistent with the complexity of the model, which in this case involves 12 latent constructs with carryover effects. Bagozzi and Yi (1994) and Fornell and Cha (1994) suggest that partial least squares (PLS) is well suited to the estimation of a complex structural equation model. It integrates aspects of principal components analysis with multiple regression (Wold 1982). When reflective measures are involved, the procedure extracts the first principal component from each subset of measures for the various latent variables and uses these principal components within a system of regression models. The algorithm then adjusts the principal component weights to maximize the predictive power of the model.

Typically, PLS models are evaluated on four key criteria: (1) the reliability of the constructs, (2) the discriminant validity of the constructs, (3) the size and significance of the path coefficients, and (4) the ability of the model to predict, in this case loyalty intentions (Hulland 1999). Regarding construct reliability, the measurement loadings should exceed .707 to ensure that at least half of the variance in the observed variable is shared with the construct. This reliability criterion is referred to as communality or, in the case of the standardized results we report here, average variance extracted (AVE; see Fornell and Larcker 1981). Table 1 reports the AVE from the PLS results for each construct for each of the three periods studied. For all 12 constructs in the model, AVE exceeds the reliability criterion. To evaluate the discriminant validity of the constructs, we examined whether the communalities or AVE measures for any two constructs that are related in the conceptual model exceed their squared correlations (Fornell and Larcker 1981). This

TABLE 1
Model Constructs, Survey Measures, and AVE by Time Period

Construct	Survey Measures	AVE
Perceived value	1. The cell phone is a good level of performance for the money I pay. 2. The cell phone is a good deal relative to other offers available in the market. 3. The price of my cell phone is more than fair for the performance I receive. 4. The cell phone is a great value.	t_1 = .686 t_2 = .741 t_3 = .649
Brand equity	1. The brand reflects my personal lifestyle. 2. The brand and my personality fit. 3. I can identify with the brand. 4. If the brand were a person, I would like to take him or her out for dinner. 5. I would like to wear clothing with the logo of my cell phone brand on it.	t_1 = .688 t_2 = .655 t_3 = .594
Affective commitment	1. I want to continue my relationship with the cell phone manufacturer. 2. The cell phone manufacturer is interested in how I use my cell phone. 3. If the cell phone manufacturer were a person, I would like to have him or her as a friend. 4. I give feedback about my evaluations of the cell phone regularly. 5. Occasionally the cell phone dealer arranges events to show new products.	t_1 = .634 t_2 = .670 t_3 = .607
Loyalty intentions	1. Next time I will definitely buy this cell phone (or its successor) again. 2. If I lose my cell phone I will definitely buy it again. 3. If I got any cell phone for free, I would choose my cell phone. 4. I recommend my cell phone to other people. 5. I talk to other people about my cell phone.	t_1 = .650 t_2 = .640 t_3 = .548

condition is satisfied in every case (not shown); thus, the discriminant validity of the constructs is also supported.

Results

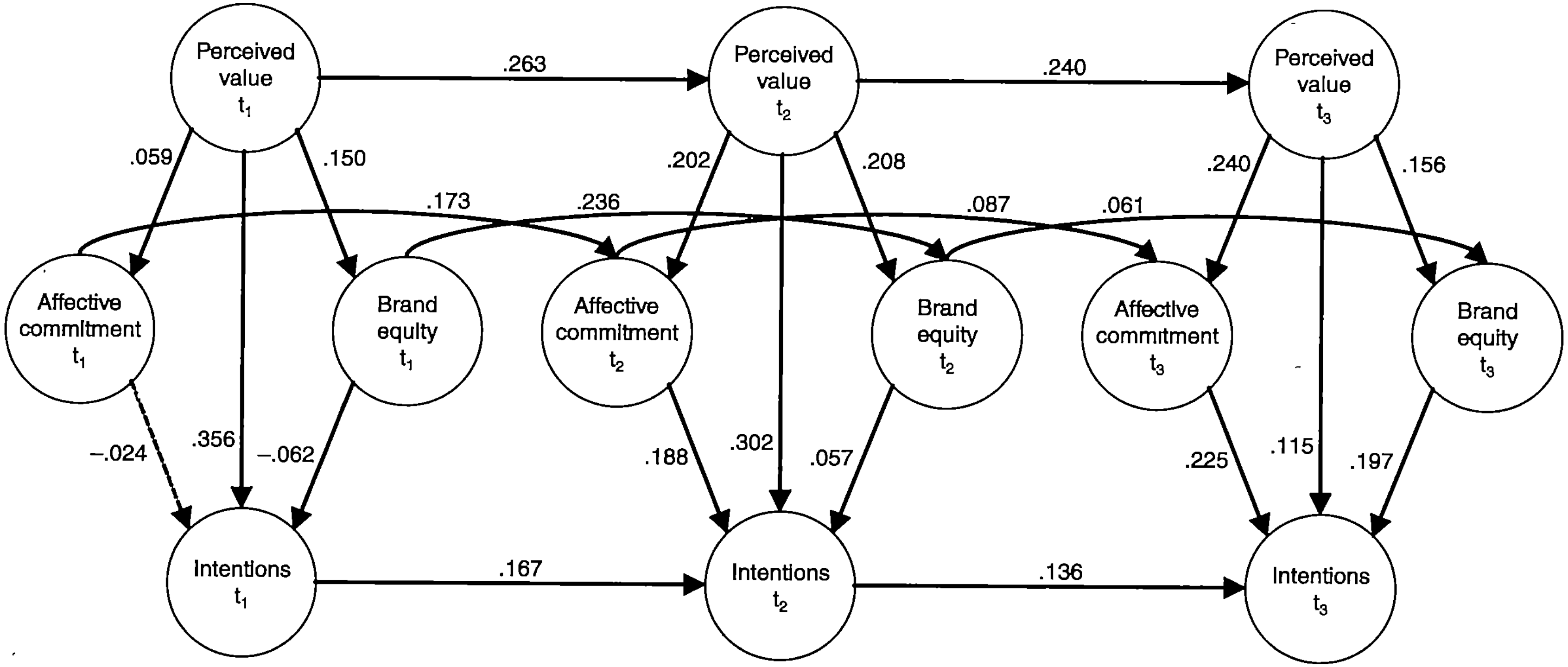
The path coefficients for the overall model appear in Figure 3. To test the statistical significance of the paths and prepare for subsequent formal tests of mediation, we used the measurement weights from the PLS model to operationalize the latent variables and rerun the final regression models. These models completely replicate the PLS results. We use the standard errors from these regressions to test for significance and significant differences in the paths over time. The only nonsignificant path in the overall model is the effect of affective commitment in t_1 on loyalty intentions in t_1. Given the nature of the hypotheses, the model results are best understood by first examining the significance of paths involving constructs in t_1 and then determining whether the effects in the model change significantly from period to period. These results appear in Table 2.

The results for t_1 (early in the diffusion of the cellular phones) show how perceived value has the only significant, positive effect on loyalty intentions. Although this perceived value makes some positive contribution to both affective commitment and brand equity, neither of these constructs makes a *positive* contribution to loyalty intentions. The unexpected results in this first period are that both affective commitment and brand equity have small negative effects on intentions, such that the negative effect of brand equity is significant. Consistent with Mittal, Kumar, and Tsiros's (1999) consumption-systems approach, all four carryover effects from t_1 to t_2 are significant (see Figure 3 and Table 2). The largest carryover effect is for perceived value (.263), followed by brand equity (.236), affective commitment (.173), and intentions (.167).

Moving from t_1 to t_2, the effect of perceived value on both affective commitment and brand equity increases. Important to the mediation argument, there is a significant decrease in the direct effect of value on intentions (from .356 to .302) combined with a significant increase in the effect of affective commitment on intentions (from –.024 to .188) and brand equity on intentions (from –.062 to .057; see Table 2). Consistent with the mediation argument, by t_2, the more affective attitudes in the model begin to mediate the effects of perceived value on intentions. Again, the carryover effects from t_2 to t_3 are positive and significant, but note the changes that occur. The carryover effects for affective commitment (.087) and brand equity (.061) from t_2 to t_3 are both significantly lower than the same effects from t_1 to t_2. This suggests that both the commitment and the image constructs are evolving as more competitors enter the market.

By t_3, perceived value continues to build relationship commitment and brand image, though the impact on brands is lower from t_2 to t_3. Importantly, the direct effect of perceived value on intentions continues to drop significantly (from .302 to .115), whereas the effects of both affective commitment and brand equity increase (for commitment, from .188 to .225; for brand, from .057 to .197). By t_3, the effects of affective commitment and brand equity on intentions are each greater than the direct effect of perceived value on intentions. Again, this is consistent with the prediction that more affective attitudes come to mediate the effects of perceived value on intentions as the market evolves. The R^2 for loyalty intentions is .124 for t_1, .201 for t_2, and .171 for t_3. There are several reasons that the vari-

FIGURE 3
PLS Path Coefficients



Notes. The dotted line represents a nonsignificant ($p < .05$) path coefficient

TABLE 2
Changes In Path Coefficients

Path Coefficient	t_1	Significant In t_1	Change from t_1 to t_2	Significant Change? (t_1 to t_2)	Change from t_2 to t_3	Significant Change? (t_2 to t_3)
Value → commitment	.059	Yes	.143	Yes	.038	No
Value → brand	.150	Yes	.058	Yes	–.052	Yes
Value → intentions	.356	Yes	–.054	Yes	–.187	Yes
Commitment → intentions	–.024	No	.212	Yes	.037	Yes
Brand → intentions	–.062	Yes	.119	Yes	.140	Yes
Current value → future value	.263	Yes	–.023	No	—	—
Current commitment → future commitment	.173	Yes	–.086	Yes	—	—
Current brand → future brand	.236	Yes	–.175	Yes	—	—
Current intentions → future intentions	.167	Yes	–.031	No	—	—

Notes: Significance is based on $p < .05$.

ance explained in loyalty intention is relatively low. This is a volatile period in the market's evolution, and we examine only the hardware side of the product/service bundle. We also expect that the variance explained increases as the market matures and buyers become loyal to particular phones and network providers.

Additional Analyses and Tests

We conducted three additional tests on the latent variable scores. First, we examined whether there were any systematic changes in the levels of the perceived value, affective commitment, brand equity, or intention constructs using a series of analysis of variance models and time as a three-level factor. Considering the large sample size involved, the differences were minimal. There were no significant differences from year to year for perceived value, affective commitment, or loyalty intentions (cumulative averages of 4.272, 4.231, and 4.257, respectively). The only difference from year to year was for brand equity, which increased significantly from t_1 (4.081) and t_2 (4.125) to t_3 (4.252; $F = 6.908$, $p < .001$). This is consistent with the increased impact of brand equity in the PLS results. Otherwise, the means are stable from period to period.

Second, we examined whether the relationships involving intentions were nonlinear, which could affect the interpretation of our model results. As a check, we used the latent variable scores to estimate linear and possible quadratic and cubic relationships between each of the three drivers of loyalty (perceived value, affective commitment, and brand equity) and loyalty intentions. In all three cases, only the linear relationship is significant; all the relationships in the model are essentially linear. An explanation for this finding is that our study focuses on the early stages of loyalty formation. We suspect that nonlinearities are more likely to occur in more mature markets in which higher levels of trust and commitment have evolved (Mittal and Kamakura 2001).

Third, we used Baron and Kenny's (1986) analysis to test for mediation among our latent constructs. With respect to our conceptual model, there are six separate sets of regression models used to test for mediation. For each year, we estimate the effect of perceived value on affective commitment (brand equity), the effect of perceived value on intentions, and the effect of affective commitment (brand equity) on intentions. Assuming that all relationships are significant, we then regress both perceived value and affective commitment (brand equity) together on intentions. Mediation is supported if the effect of value on intentions drops significantly. Given the number of equations involved, we summarize the test results here. In t_1, there is no support for mediation involving either affective commitment or brand equity. In t_2, the tests support partial mediation for the relationship from perceived value, to affective commitment, to intentions but not for the relationship from value, to brand equity, to intentions. By t_3, the tests support partial mediation for both sets of relationships. In general, these results are consistent with the PLS analysis.

Results Summary

We summarize our findings in terms of our hypotheses. In support of H_1, we find that perceived value has a positive direct effect on loyalty intentions that decreases over time. In support of H_2, value has positive effects on affective commitment that grow over time. The support for H_3 is mixed. Although the positive effect of value on brand equity grows from t_1 to t_2, it declines again in t_3. In each year, however, the effect is significantly positive. Affective commitment has no significant effect on intentions in t_1, but the effect becomes progressively positive over time. Likewise, brand equity has a small negative effect on intention in t_1, which becomes progressively positive over time. Thus, H_4 and H_5 are partially supported. When we combine the analysis of path coefficients with the formal tests of mediation, there is robust support for the prediction that intentions are predominantly value driven early in a life cycle and are increasingly mediated by attitudes toward the brand and relationship over time. Finally, consistent with Mittal, Kumar, and Tsiros's (1999) consumption-system framework, our results support significant carryover effects from year to year for all our latent variables.

Discussion and Implications

Recent dynamic models in marketing confirm that the drivers of loyalty intentions evolve over time. Yet these studies are limited to relatively mature product and service categories. Our study provides two important contributions to the understanding of the evolution of customer evaluations, attitudes, and intentions. First, we demonstrate how the drivers of intentions evolve for a panel of loyal customers who purchased and used a new-to-market offering, cellular phone technology, through the introduction and growth phases of a life cycle. Second, we demonstrate how two constructs that have not been included in previous studies, affective commitment and brand equity, come to mediate the effects of perceived value on intentions over time.

We find that cognitive perceptions of overall value drive loyalty intentions early in the diffusion process. As the market grows and customer experience accumulates, more affective attitudes toward the relationship and brand come to drive intentions. Our predictions and findings are consistent with attitude theory; performance beliefs have more direct effects on behavioral intentions when experience is low, whereas attitudes come to mediate the effects of performance beliefs on intentions as experience increases. An important theoretical implication is that affective commitment and brand equity are effective ways to operationalize the affect-based mediators of perceived value on intentions.

Our findings both replicate and extend our understanding of consumption systems (Mittal, Kumar and Tsiros 1999). Consistent with previous studies, we find that performance perceptions (perceived value) and intentions in one period serve as anchors for the same perceptions and intentions in subsequent periods. Building on previous studies, we introduce two new constructs to the consumption system, affective commitment and brand equity. One important finding is that both the commitment and the brand equity constructs exhibit significant carryover effects. Another finding is that, when we control for carryover effects, affective commitment and brand equity come to mediate the effects of perceived performance on intentions over time.

The results have important implications for managers who are responsible for improving perceived value and managers who are responsible for building brands and relationships. Early in the introduction and growth of an offering, managers should maintain a focus on improving value. With a focus on value creation, a foundation is laid for building brands and relationships. As a market grows, both brands and relationships become more important. Notably, brands and relationships had small negative impacts on loyalty intentions in the first phase of our study, perhaps because our sample comprised technology upgraders or early adopters. As the market grew to include a variety of competitors, the influence of brands and relationships on intentions became positive and eventually exceeded the direct, positive influence of perceived value.

With respect to relationship commitment, the implication is that the personal relationship between a customer and a company should be measured, nurtured, and managed effectively (Bendapudi and Leone 2002) through a company's customer relationship management system. For brand equity, the implication is that the degree to which customers identify with a particular brand and find it relevant to their situation should also be measured and managed effectively. Our results provide insight into how much time it takes for brands and relationships to influence intentions. Within a period of two years, affective commitment gains a sizable influence on intentions. Within a period of four years, both commitment and brand equity have direct effects on intentions that exceed the direct effect of perceived value. Note, however, that perceived value continues to influence loyalty through the mediating effects of the brand and relationship.

In other words, managers must recognize how life inside a cellular phone or the technology per se is replaced by life outside the cellular phone over time. Early in the diffusion process, technology determines the product policy of cellular phone manufacturers. As the technology becomes accepted, other forms of differentiation, such as colors and design, become increasingly important. Parallel to this, whether these designs, colors, and associated brand names fit a customer's own personality become critical. Adaptation to this evolution requires shifting from a focus on how a technology serves a need to a focus on other types of marketing expertise, specifically the building of brands and relationships. Several factors should affect this shift. The first factor is customer involvement; the more customers are involved in a particular product category, the more they seek brand and relationship benefits when the product benefits become standard. The second factor is competition; more competitive markets have more players that attempt to differentiate themselves. The third factor is core product standardization; the more inherently homogeneous the product, the more companies must focus on brands and relationships as a source of differentiation.

The most important contribution of this study is that it improves the basic understanding of how perceived value, brands, and relationships come to influence one another and loyalty intentions over the introduction and growth phases of a life cycle. A limitation of our study is that we focus on the cellular phone rather than the entire product/service bundle. An important reason for this is that the survey is from the perspective of the cellular phone manufacturer in an environment in which cellular phone brands and service providers are interchangeable. Another limitation is that our panel is limited to consumers who enter a market early and remain loyal to a particular brand. The results should also be replicated in other contexts before generalizations can emerge. It will be particularly important to follow the category forward as it reaches maturity to help understand how the drivers of both intentions and behavior stabilize. Finally, our measures of relationship commitment focus on affective commitment. Because the market for cellular phones includes significant switching barriers, it is important to include calculative commitment when further testing the model.

REFERENCES

Aaker, David A. (2004), *Brand Portfolio Strategy: Creating Relevance, Differentiation, Energy, Leverage, and Clarity*. New York: The Free Press.

Anderson, Erin and Barton Weitz (1992), "The Use of Pledges to Build and Sustain Commitment in Distribution Channels," *Journal of Marketing Research*, 29 (February), 18–34.

Bagozzi, Richard P. (1975), "Marketing as Exchange," *Journal of Marketing*, 39 (October), 32–39.

——— and Paul R. Warshaw (1990), "Trying to Consume," *Journal of Consumer Research*, 17 (September), 127–40

——— and Youjae Yi (1994), "Advanced Topics in Structural Equation Models," in *Advanced Methods of Marketing Research*, Richard P. Bagozzi, ed. Cambridge, MA. Blackwell, 1–52.

Baron, Reuben M. and David A. Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research," *Journal of Personality and Social Psychology*, 51 (6), 1173–82.

Bendapudi, Neeli and Leonard L. Berry (1997), "Customers' Motivations for Maintaining Relationships with Service Providers," *Journal of Retailing*, 73 (1), 15–37

——— and Robert P. Leone (2002), "Managing Business-to-Business Customer Relationships Following Key Contact Employee Turnover in a Vendor Firm," *Journal of Marketing*, 66 (April), 83–101.

Berry, Leonard L. and A. Parasuraman (1991), *Marketing Services*. New York: The Free Press.

Bolton, Ruth N. (1998), "A Dynamic Model of Customers' Usage of Services: Usage as an Antecedent and Consequence of Satisfaction," *Marketing Science*, 17 (Winter), 45–65.

——— and James H. Drew (1991), "A Longitudinal Analysis of the Impact of Service Changes on Customer Attitudes," *Journal of Marketing*, 55 (January), 1–9.

——— and Katherine N. Lemon (1999), "A Dynamic Model of Customers' Usage of Services: Usage as an Antecedent and Consequence of Satisfaction," *Journal of Marketing Research*, 36 (May), 171–86.

Dwyer, F. Robert, Paul H. Schurr, and Sejo Oh (1987), "Developing Buyer–Seller Relationships," *Journal of Marketing*, 51 (April), 11–27.

Fazio, Russell H., Martha C. Powell, and Carol J. Williams (1989), "The Role of Attitude Accessibility in the Attitude-to-Behavior Relationship," *Journal of Consumer Research*, 16 (December), 280–88.

Fornell, Claes and Jaesung Cha (1994), "Partial Least Squares," in *Advanced Methods of Marketing Research*, Richard P. Bagozzi, ed. Cambridge, MA: Blackwell, 52–78.

———, Michael D. Johnson, Eugene W. Anderson, Jaesung Cha, and Barbara Everitt Bryant (1996), "The American Customer Satisfaction Index: Nature, Purpose, and Findings," *Journal of Marketing*, 60 (October), 7–18.

——— and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39–50.

Fullerton, Gordon (2003), "When Does Commitment Lead to Loyalty?" *Journal of Service Research*, 5 (May), 333–44.

Garbarino, Ellen and Mark S. Johnson (1999), "The Differential Roles of Satisfaction, Trust, and Commitment in Customer Relationships," *Journal of Marketing*, 63 (April), 70–87.

Gruen, Thomas, John Summers, and Frank Acito (2000), "Customer-Employee Rapport in Service Relationships," *Journal of Service Research*, 3 (August), 34–49.

Gustafsson, Anders, Michael D. Johnson, and Inger Roos (2005), "The Effects of Customer Satisfaction, Relationship Commitment Dimensions, and Triggers on Customer Retention," *Journal of Marketing*, 69 (October), 210–18.

Hansen, Håvard, Kåre Sandvik, and Fred Selnes (2003), "Direct and Indirect Effects of Commitment to a Service Employee on the Intention to Stay," *Journal of Service Research*, 5 (May), 356–68.

Heilman, Carrie M., Douglas Bowman, and Gordon P. Wright (2000), "The Evolution of Brand Preference and Choice Behaviors of Consumers New to a Market," *Journal of Marketing Research*, 37 (May), 139–55.

Helson, Harry (1948), "Adaptation-Level as a Basis for a Quantitative Theory of Frames of Reference," *Psychological Review*, 55 (November), 297–313.

Hulland, John (1999), "Use of Partial Least Squares (PLS) in Strategic Management Research: A Review of Four Recent Studies," *Strategic Management Journal*, 20 (February), 195–204.

Johnson, Michael D., Anders Gustafsson, Tor Wallin Andreassen, Line Lervik, and Jaesung Cha (2001), "The Evolution and Future of National Customer Satisfaction Index Models," *Journal of Economic Psychology*, 22 (April), 217–45.

Keller, Kevin Lane (1993), "Conceptualizing, Measuring, and Managing Customer-Based Brand Equity," *Journal of Marketing*, 57 (January), 1–22.

——— (2003), "Brand Synthesis: The Multidimensionality of Brand Knowledge," *Journal of Consumer Research*, 29 (March), 595–600.

Mittal, Vikas and Wagner Kamakura (2001), "Satisfaction, Repurchase Intent, and Repurchase Behavior: Investigating the Moderating Effects of Customer Characteristics," *Journal of Marketing Research*, 38 (February), 131–42.

———, Jerome M. Katrichis, and Pankaj Kumar (2001), "Attribute Performance and Customer Satisfaction over Time: Evidence from Two Field Studies," *Journal of Services Marketing*, 15 (5), 343–56.

———, Pankaj Kumar, and Michael Tsiros (1999), "Attribute-Level Performance, Satisfaction, and Behavioral Intentions over Time: A Consumption-System Approach," *Journal of Marketing*, 63 (April), 88–101.

Morgan, Robert M. and Shelby D. Hunt (1994), "The Commitment–Trust Theory of Relationship Marketing," *Journal of Marketing*, 58 (July), 20–38.

Oliver, Richard L. (1980), "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research*, 17 (November), 460–69.

——— (1999), "Whence Consumer Loyalty?" *Journal of Marketing*, 63 (Special Issue), 33–44.

Olsen, Line Lervik and Michael D. Johnson (2003), "Service Equity, Satisfaction, and Loyalty. From Transaction-Specific to Cumulative Evaluations," *Journal of Service Research*, 5 (February), 184–95.

Parasuraman, A. (2000), "Technology Readiness Index (TRI): A Multiple-Item Scale to Measure Readiness to Embrace New Technologies," *Journal of Service Research*, 2 (May), 307–320.

Priester, Joseph R., Dhananjay Nayakankuppam, Monique A. Fleming, and John Godek (2004), "The A^{2}SC^2 Model: The Influence of Attitudes and Attitude Strength on Consideration and Choice," *Journal of Consumer Research*, 30 (March), 574–87.

Rosen, Emanuel (2000), *The Anatomy of Buzz: How to Create Word of Mouth Marketing*. New York: Doubleday.

Rust, Roland T., Katherine N. Lemon, and Valarie A. Zeithaml (2004), "Return on Marketing: Using Customer Equity to Focus Marketing Strategy," *Journal of Marketing*, 68 (January), 109–121.

———, Valarie A. Zeithaml, and Katherine N. Lemon (2000), *Driving Customer Equity: How Customer Lifetime Value Is Reshaping Corporate Strategy*. New York: The Free Press

Selnes, Fred (1993), "An Examination of the Effect of Product Performance on Brand Reputation, Satisfaction and Loyalty," *European Journal of Marketing*, 27 (9), 19–35.

Slotegraaf, Rebecca J. and J. Jeffrey Inman (2004), "Longitudinal Shifts in the Drivers of Satisfaction with Product Quality: The Role of Attribute Resolvability," *Journal of Marketing Research*, 41 (August), 269–80.

Verhoef, Peter C. (2003), "Understanding the Effect of Customer Relationship Management Efforts on Customer Retention and Customer Share Development," *Journal of Marketing*, 67 (October), 30–45.

Wold, Herman (1982), "Systems Under Indirect Observation Using PLS," in *A Second Generation of Multivariate Analysis Methods*, Claes Fornell, ed. New York: Praeger.

Zeithaml, Valarie A., Leonard L. Berry, and A. Parasuraman (1996), "The Behavioral Consequences of Service Loyalty," *Journal of Marketing*, 60 (April), 31–46.

Ann E. Schlosser, Tiffany Barnett White, & Susan M. Lloyd

Converting Web Site Visitors into Buyers: How Web Site Investment Increases Consumer Trusting Beliefs and Online Purchase Intentions

The authors investigate the impact of Web site design investments on consumers' trusting beliefs and online purchase intentions. Such investments signal the component of trusting beliefs that is most strongly related to online purchase intentions: ability. These effects were strongest when consumers' goals were to search rather than to browse and when purchases involved risk.

"In important ways, using the Internet involves a leap of faith. We type in our credit card numbers and other personal information in order to make purchases over the Internet and trust that this information will not be used in unauthorized or fraudulent ways" (Bargh and McKenna 2004, p. 585). Firms have responded to such consumer concerns by investing in Web site security, which has become a multibillion dollar industry (eMarketer 2005). Yet even experienced online buyers view purchasing online as risky (eMarketer.com 2005; Forrester 2005). Indeed, consumers' perceptions of the risks involved in providing personal information online often contrast the views of security experts, causing these consumers to avoid online activities that are actually safe (Dunn 2004). This avoidance has led some experts to speculate that the immediate threat to e-commerce is consumers' perceptions (eMarketer 2005; Hoffman, Novak, and Peralta 1999; Rust, Kannan, and Peng 2002). Thus, although it is necessary, investing in back-end technology to protect consumers' information and ensure that e-commerce transactions run smoothly is not sufficient. Firms, particularly those attempting to convert visitors to buyers, still face the challenge of establishing consumers' trust online. Thus, it is important to understand how companies can communicate their trustworthiness to consumers.

Marketing managers face the challenge of establishing consumers' trust in a variety of contexts, but doing so in computer-mediated environments such as the Internet may be particularly difficult (Naquin and Paulson 2003). A common approach is to post explicit statements that assure customers that personal data will be discreetly used and protected (i.e., privacy and security statements). However, evidence on the effectiveness of such statements is contradictory. Whereas some research has shown that such statements help instill consumer confidence in e-commerce sites (Palmer, Bailey, and Faraj 2000), others suggest that they are not necessarily the most important predictor of online trust (Montoya-Weiss, Voss, and Grewal 2003; Sultan et al. 2002). Findings from a recent large-scale study suggest that despite consumers' claims that privacy policies are important for establishing credibility, consumers refer instead to "surface" elements, such as Web site design (Fogg et al. 2002).

In this article, we develop a conceptual framework for understanding how marketing signals influence consumers' trust in an e-commerce setting. We clarify important distinctions related to trust that have been largely overlooked in the literature but are key to understanding how online firms can best convert visitors to buyers. Specifically, we argue that different Web site signals can influence different beliefs about a firm's trustworthiness, which in turn have differential effects on online purchase intentions. Furthermore, these effects might vary according to the consumers' purpose for visiting the site and the level of risk they perceive in the purchase. Consistent with prior research (e.g., McKnight, Choudhury, and Kacmar 2002), we conceptualize online purchase intentions in terms of customer acquisition—that is, consumers' intentions to make an initial online purchase from a firm, despite their online purchase history with other firms. We begin with a review of the trust literature. We then present and test our conceptual framework and conclude with a discussion of the theoretical, managerial, and public policy implications of these findings.

Ann E. Schlosser is Assistant Professor of Marketing, University of Washington Business School (e-mail: aschloss@u.washington.edu). Tiffany Barnett White is Assistant Professor of Business Administration (Marketing Group), College of Business, University of Illinois (e-mail: tbwhite@uiuc.edu). Susan M. Lloyd is Assistant Professor of Marketing and Endowed Fellow, Kogod School of Business, American University (e-mail: smlloyd@american.edu). The authors thank Ruth Bolton and the two anonymous *JM* reviewers, as well as Jim Bettman, Katherine Lemon, Michael Mazis, and Kent Monroe, for their helpful comments on previous versions of this article. The authors also thank Sandra M. Dahne and S. Rand Fishkin for their technical assistance. Funding for Study 1 was provided by the Center for Information Technology and the Global Economy at the Kogod School of Business, American University.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 70 (April 2006), 133–148

Trust

Trust has been defined as "a willingness to rely on an exchange partner in whom one has confidence" (Moorman, Zaltman, and Deshpandé 1992, p. 315); "a generalized expectancy held by an individual that the word, promise, oral or written statement of another individual or group can be relied upon" (Rotter 1980, p. 1); and "a belief in a person's competence to perform a specific task under specific circumstances" (Sitkin and Roth 1993, p. 373). Reflected in these and other definitions of trust is a cognitive aspect (i.e., trusting beliefs) and a behavioral aspect (i.e., trusting intentions) (Kim et al. 2004; Moorman, Zaltman, and Deshpandé 1992).

Trusting beliefs represent a "sentiment, or expectation about an exchange partner's trustworthiness" (Moorman, Deshpandé, and Zaltman 1993, p. 315). Although various trusting beliefs have been studied in the literature, the majority can be conceptually clustered into three dimensions: ability, benevolence, and integrity (McKnight, Choudhury, and Kacmar 2002). "Ability beliefs" reflect consumers' confidence that the firm has the skills necessary to perform the job (Mayer, Davis, and Schoorman 1995), "benevolence beliefs" reflect confidence that the firm has a positive orientation toward its consumers beyond an "egocentric profit motive" (Mayer, Davis, and Schoorman 1995, p. 717), and "integrity beliefs" reflect confidence that the firm adheres to a set of moral principles or professional standards that guide its interactions with customers. These trusting beliefs are related, yet distinct. For example, consumers may believe that the firm cares about its customers and thus intends to deliver a smooth, error-free transaction (i.e., the firm is benevolent), but they may also believe that the firm lacks the ability to do so. Likewise, although integrity and benevolence beliefs are similar, the former focuses on meeting objective standards of corporate citizenship, and the latter focuses on customer welfare that goes beyond normal business activity. For example, despite consumers' beliefs that the firm follows a professional code of conduct (i.e., has integrity), they may still question the firm's genuine concern for its customers (i.e., its benevolence).

Although ability, benevolence, and integrity beliefs are acknowledged as conceptually distinct (e.g., Kumar, Scheer, and Steenkamp 1995), they are often combined into a global measure of trusting beliefs (e.g., Doney and Cannon 1997). Whereas combining these beliefs into a single variable is a parsimonious approach to studying trust, it can make it difficult to identify what action should be taken to build trust (Smith and Barclay 1997). Because a global measure likely obscures the reason certain signals are more effective than others in affecting online purchase intentions, we treat each trusting belief separately.

Trusting intentions represent "a willingness to make oneself vulnerable to another in the presence of risk" (Kim et al. 2004, p. 105). What distinguishes trusting intentions from other types of behavioral intentions is that they involve risk (Moorman, Zaltman, and Deshpandé 1992). As reflected in the opening quotation, purchasing online involves risk, especially when a person lacks experience with the online firm. Specifically, the consumer must be willing to transfer resources (e.g., credit card and other personal information) to the online firm, the consequences of which could be damaging. For example, among the real and/or perceived risks are that the firm may overcharge, fail to deliver the product, deliver an inferior product, or fail to protect personal information. To the extent that consumers are concerned about these and other risks of purchasing online, online purchase intentions reflect trusting intentions.

The distinction between trusting beliefs and trusting intentions has been acknowledged by some researchers (e.g., Moorman, Zaltman, and Deshpandé 1992; Sirdeshmukh, Singh, and Sabol 2002) but ignored by others who have studied only trusting beliefs, implicitly assuming that these beliefs imply trust (e.g., Doney and Cannon 1997; Ganeson 1994; Kumar, Scheer, and Steenkamp 1995; Mayer, Davis, and Schoorman 1995; Morgan and Hunt 1994). For example, Morgan and Hunt (1994) argue that trusting beliefs are sufficient for measuring trust because such beliefs imply that trusting intentions will follow. In contrast, Moorman, Zaltman, and Deshpandé (1992) argue that trust is limited when trusting beliefs do not accompany a corresponding trusting intention or when trusting intentions occur without corresponding trusting beliefs (e.g., under conditions of coercion or limited alternatives). In other words, these researchers argue that both trusting beliefs and trusting intentions must be present for trust to exist. Likewise, we argue that trusting beliefs are a necessary but not sufficient condition for trust to exist, because increasing trusting beliefs will not always have a corresponding positive effect on trusting intentions.

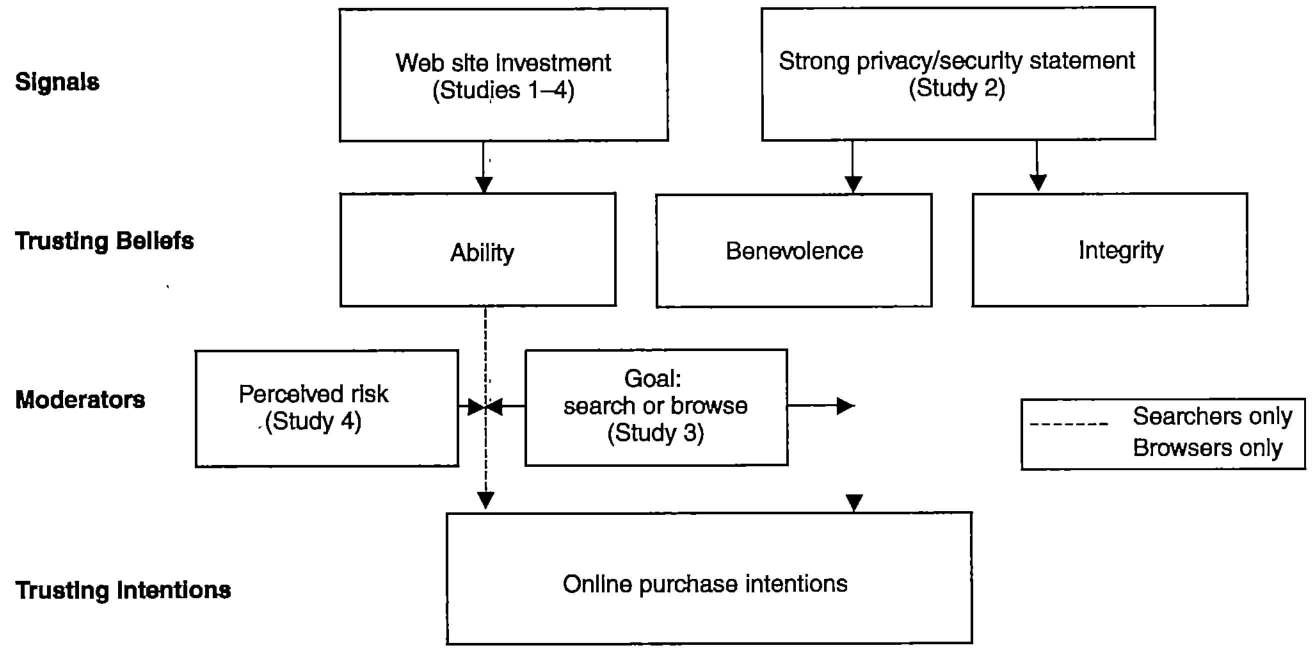
Conceptual Framework

Drawing from research on trust (Mayer, Davis, and Schoorman 1995; Moorman, Deshpandé, and Zaltman 1993), consumer goals (e.g., Hoffman and Novak 1996), and marketing signals (e.g., Kirmani and Wright 1989; Prabhu and Stewart 2001), we develop a conceptual framework for understanding how different signals influence ability, benevolence, and integrity beliefs and thus influence online purchase intentions (see Figure 1). Because our objective is to understand how to increase consumers' willingness to buy online, we focus on those whose goal is most consistent with buying online, namely, consumers who search for product information (or searchers; Hoffman and Novak 1996; Moe 2003; Schlosser 2003). Indeed, searchers think about and are persuaded more by product information (Schlosser 2003) and have higher visitor-to-buyer conversion rates than those who do not search (Moe 2003). We begin by examining the relationship between searchers' trusting beliefs and intentions.

Searchers' Trusting Beliefs and Online Purchase Intentions

Searching reflects purposive, task-specific behaviors, such as the planned acquisition of information during prepurchase deliberation (Hoffman and Novak 1996; Janiszewski 1998). Similar to those who read a text to find an answer to a question (Rosenblatt 1978), searchers are likely motivated to find the right answer efficiently. Such a fact-gathering,

FIGURE 1
Conceptual Framework of the Effect of Online Signals on Trusting Beliefs and Intentions



knowledge-seeking stance is typically outcome oriented, concentrated, impersonal, and objective (Rosenblatt 1978). Given this performance orientation, we expect that when considering whether to purchase online, searchers will focus on the trusting belief that is most relevant to performance: ability (Mayer, Davis, and Schoorman 1995). If this is the case, searchers' ability beliefs should largely influence their online purchase intentions. In contrast, their beliefs about the firm's trustworthiness on nonperformance-related dimensions (i.e., benevolence and integrity) should have relatively little effect on their online purchase intentions.

H_1: Searchers' online purchase intentions depend on their beliefs about the firm's ability rather than their beliefs about the firm's benevolence or integrity.

Signaling Ability Through Web Site Investment

If trust in a firm's abilities is critical to increasing online purchase intentions, a fundamental question is, How can firms use online cues to communicate that their abilities can be trusted? To address this question, we draw from research on marketing signals. Signals are the actions or announcements that convey a firm's abilities and intentions (Porter 1980). Marketers often use observable signals (e.g., price, warranties, advertising expenditures) to communicate the level of some unobservable quality (e.g., product quality; Kirmani and Rao 2000). Signaling may be especially important in an online purchasing context because of the inherent asymmetry of relevant information between buyers and sellers. Specifically, information about the quality of a given online transaction is generally unobservable by consumers before purchase.

Investing in Web site design may be one observable signal that firms can use to communicate their abilities and boost searchers' online purchase intentions. Indeed, consumers can distinguish between expensive and inexpensive marketing tactics, such as ad production elements (Kirmani and Wright 1989). Moreover, they make inferences about companies (e.g., the company's ability to make quality products and its credibility) based on these perceived marketing expenditures (for a review, see Kirmani and Rao 2000). The attribution literature provides insight into these findings (Kirmani and Wright 1989). Specifically, when interpreting others' performance, people infer that investing time and energy promotes success (Weiner 1986). Likewise, consumers will likely infer that a firm that has invested in Web site design can successfully handle online transactions.

As in prior research, we consider investment in broad terms (Kirmani and Wright 1989); that is, investment reflects expenditures of time, money, and effort to Web site design. Importantly, it refers to investments in the front-end (design) elements of a Web site (i.e., its observable characteristics) and not back-end technologies, such as order fulfillment software, security encryption, and firewall capabilities, which are typically unobservable before purchase. Yet because ability is a stable, internal characteristic (Weiner 1972), people will likely generalize their trust in a firm's ability in one area (design) to other related areas (e.g., order fulfillment). Thus, instead of being purely cosmetic, Web site design likely communicates important performance information. However, it likely communicates less about the firm's goodwill, ethics, values, or intentions to mislead than it does about the firm's ability.[1] Consequently, consumers will likely use perceived Web site investment to infer a firm's ability more than its benevolence or integrity. If this is the case and ability beliefs largely influence searchers' online purchase intentions (H_1), it follows that

---

[1] This assumes conditions of normality. Because a poorly designed Web site (e.g., one with hyperlinks that do not work)

H_2: Searchers' online purchase intentions will be higher at a high-investment Web site than at a low-investment Web site.

H_3: Beliefs about the firm's ability will mediate the relationship between site investment and searchers' online purchase intentions.

The first two studies test these hypotheses, and the last two studies examine two potential moderators (goal and perceived risk). Furthermore, to test the generalizability of these effects, we varied across studies the samples recruited (students versus nonstudents), companies (a fictitious versus well-known firm), and products (home furnishings and accessories versus cameras).

Study 1

Method

Sample and design. The sample consisted of 111 respondents who participated in exchange for $10 and were recruited through an electronic and a printed newsletter distributed to university employees. The sample was 68% female, with a mean age of 37.5 and a mean income of $35,000 to $49,999. Respondents had a median education of four years of college and used the Internet an average of four to six times per week. We randomly assigned respondents to a high- or low-investment site.

Web site investment manipulation. We manipulated Web site investment through the presence of sophisticated Web site technology and visual design elements. Specifically, the high-investment site had a white background, sophisticated fonts (images for the navigation bar; Garamond font), and an enhanced zoom feature created with Design Within Reach. This enabled users to zoom in on any part of an image and to choose among three preselected zooms that executed automatically with a single click. In contrast, the low-investment site featured the default background color and font (gray; Times New Roman) and a limited zoom feature, which, when clicked, simply provided a larger view of the focal product. The content and layout of both sites were identical.

To test the effectiveness of this manipulation, 43 undergraduates viewed screen captures of the homepage (font and background color were used to manipulate investment) and a zoom page (technology was used to manipulate investment). The order was counterbalanced, and participants viewed only those pages specific to the high- or low-investment site. After viewing each page, participants reported how much time, effort, and money they believed the firm invested in each page on a scale from 1 ("very little") to 7 ("a great deal"). We averaged the responses to these items ($\alpha s > .94$) and analyzed them with a 2 (investment: high versus low) × 2 (page: zoom versus home) × 2 (order: viewed home or zoom page first) analysis of variance (ANOVA). In support of the investment manipulation, participants perceived the high-investment site as requiring greater investment than the low-investment site (Ms = 4.09 versus 3.18; $F(1, 39) = 9.02$, $p < .01$). Furthermore, the effectiveness of the investment manipulation was unaffected by the type or order of page viewed ($Fs(1, 39) < 3.35$, not significant [n.s.]). Thus, it appears that font, background color, and use of technology communicate investment. Participants also reported how informative, entertaining, and well organized they found the pages to be on the same seven-point scale. As we expected, investment did not affect these variables ($Fs(1, 39) < 1.97$, n.s.).

Procedure. Participants sat at a computer and received a paper booklet. The first page instructed participants that they would be visiting a site for a new firm called UrbanFurniture (UF) and to limit their visit to the homepage and living-room sections of the site. All participants were told to imagine that they wanted to purchase contemporary furniture and were considering UF. As in prior research (Schlosser 2003), to instill a searching goal, participants were asked to write down two questions they had for UF about its products before visiting the site.

Participants then visited either the high- or the low-investment site, which was preloaded on their computer. Afterward, they completed the survey, which contained three items that measured their online purchase intentions ($\alpha = .91$) and a modification of Mayer and Davis's (1999) scale of trustworthiness (for the measures used in this and the other studies, see the Appendix). This scale measured beliefs about UF's ability ($\alpha = .90$), benevolence ($\alpha = .88$), and integrity ($\alpha = .71$). Then, to test the effectiveness of the investment manipulation in the main experiment, participants completed the three-item Web site investment scale (see the Appendix; $\alpha = .95$).

At the end of the survey, participants reported their education and income levels as well as how often they used the Internet to purchase goods in the last six months on a scale ranging from 1 ("not at all") to 7 ("quite often"). We included these measures to control for individual differences in Internet experience both directly (i.e., through self-reported use of the Web) and indirectly (i.e., using demographic variables associated with Web use). To control for variance due to mechanical error, we asked participants to report the extent to which they encountered problems at the UF site (e.g., error messages, server delays, crashing) on a scale from 1 ("not at all") to 7 ("quite often").

Results

Manipulation checks. We analyzed the investment manipulation with a one-way analysis of covariance (ANCOVA), controlling for reported problems with the site, prior Web purchase history, education, and income. In support of the manipulation, perceived Web site investment was higher among those in the high-investment condition than among those in the low-investment condition (Ms = 3.46 versus 1.97; $F(1, 108) = 32.45$, $p < .0001$).

Trusting beliefs and online purchase intentions. We used hierarchical regression to test H_1. We first modeled

---

likely violates consumers' expectations about norms of conduct for e-commerce firms, it will likely raise concerns about the firm's benevolence and integrity.

online purchase intentions as a function of ability beliefs. As we predicted, ability significantly influenced these intentions ($\beta = .27$, $t(108) = 2.88$, $p < .005$; $R^2 = .07$, $F(1, 108) = 8.31$, $p < .01$; see Table 1). Furthermore, the addition of benevolence and integrity beliefs to the model did not contribute significantly to explaining searchers' online purchase intentions ($\Delta R^2 = .007$, $F(2, 106) < 1$). We replicated this pattern of results with a stepwise regression analysis, which identifies the best subset of belief variables that predict online purchase intentions. Thus, regardless of the regression procedures we used, the results are consistent with H_1. Furthermore, given the inherent correlation between trusting beliefs, we tested for multicollinearity by examining the maximum variance inflation factor (VIF). Multicollinearity is problematic for interpreting regression analyses when the maximum VIF is greater than 10 (Neter, Wasserman, and Kutner 1990). We found that multicollinearity was not an issue in the preceding analysis (the maximum VIF = 3.10).

Web site investment, trusting beliefs, and online purchase intentions. We analyzed ability beliefs with a one-way ANCOVA. As we expected, ability beliefs were higher for those who visited the high-investment site than for those who visited the low-investment site (Ms = 2.41 versus 1.80; $F(1, 109) = 13.27$, $p < .005$). In contrast, Web site investment did not affect benevolence or integrity beliefs ($Fs(1, 109) \leq 2.00$, n.s.). Thus, Web site investment is more effective in communicating the trustworthiness of a firm's ability than its benevolence and integrity.

In support of H_2, online purchase intentions were higher among those who visited the high-investment site than among those who visited the low-investment site (Ms = –.13 versus –.78; $F(1, 109) = 4.06$, $p < .05$). To test whether ability beliefs mediated this effect (H_3), we added ability beliefs as a covariate to the ANCOVA. Consistent with the requirements for mediation (Baron and Kenny 1986), ability was significant ($F(1, 108) = 4.31$, $p < .05$), and the investment effect became nonsignificant ($F(1, 108) = 1.48$, n.s.). We found further support for H_3 using the criteria that Sobel (1982) endorsed for testing mediation (Goodman I test statistic = 2.22; $p < .05$). Because investment did not affect benevolence and integrity beliefs, these beliefs cannot be considered mediators.

Conclusions

Study 1 provides support for H_1–H_3. Specifically, searchers' online purchase intentions were influenced by their trust in the firm's ability rather than their trust in its benevolence and integrity. As a result, ability signals (i.e., Web site investment) influenced their online purchase intentions. Furthermore, their trust in the firm's ability mediated this effect.

It is possible that searchers' online purchase intentions were influenced by their ability beliefs rather than by their benevolence and integrity beliefs because there were no signals regarding the firm's benevolence and integrity. In the presence of such signals, ability beliefs may have less influence. Indeed, the impact of a given signal is weakened when other, more relevant signals are present (Kirmani and Wright 1989). If benevolence and integrity signals are more relevant, the effect of investment on searchers' online purchase intentions should be weaker when such signals are present. However, if ability is more relevant, investment should affect searchers' online purchase intentions regardless of whether benevolence and integrity signals are present. We directly test this in Study 2 by manipulating the presence and strength of a firm's privacy and security statement. Because benevolence represents the firm's orientation toward customers and integrity represents whether the firm will do what it promises (Mayer, Davis, and Schoorman 1995), one method of signaling a firm's benevolence and integrity may be through formal statements of its intentions to consumers, such as through privacy and security statements. If so, such statements should affect consumers' beliefs about a firm's benevolence and integrity rather than its ability. Yet if ability is a stronger driver of searchers'

TABLE 1
Hierarchical Regression Analysis of Trusting Beliefs on Online Purchase Intentions

			Model 1			Model 2		
Study	**Goal**	**Variable**	**β**	**R^2**	**F for R^2**	**β**	**ΔR^2**	**F for ΔR^2**
1	Searchers	Ability	.27*	.07	8.31*	.19	.01	.38
		Benevolence				.10		
		Integrity				.02		
2	Searchers	Ability	.39*	.15	15.57*	.39*	.00	.08
		Benevolence				.06		
		Integrity				–.06		
3	Searchers	Ability	.42*	.18	15.70*	.43*	.03	1.50
		Benevolence				.15		
		Integrity				–.20		
	Browsers	Ability	.15	.02	1.62	.10	.10	4.28*
		Benevolence				.33*		
		Integrity				–.01		

*$p \leq 05$.

online purchase intentions (H_1), investment should affect searchers' online purchase intentions regardless of whether a privacy and security statement is present. That is, H_2 and H_3 should be supported even when such a statement is provided.

Study 2

Method

Sample and design. A total of 79 undergraduate students participated in exchange for extra course credit. We randomly assigned participants to one of six conditions in a 2 (investment) × 3 (privacy/security statement: strong or weak versus absent) design.[2] We included the absent condition to examine whether the presence of a strong statement is better than having no statement. We included the weak condition to examine whether merely having a privacy/security statement might signal benevolence and integrity or whether the contents of the statement influence such beliefs.

Privacy/security statement manipulation. We constructed the strong and weak privacy/security statements on the basis of a content analysis of the privacy/security statements gathered from more than 25 sites. For the strong statement, explicit information was available about how UF collects and uses customer information. There was also a promise of confidentiality, a contact number, an opt-in feature, encryption information, and a 100% guarantee against information theft. In contrast, the weak statement informed the consumer that personal information would be collected and made available to other vendors that "are offering products we feel are of interest to you." There was no opportunity to opt in or out of such correspondence. Consumers were also informed that UF "tries to safely transmit your account information." No account protection or guarantee was offered.

To pretest this manipulation, 37 undergraduates read either the strong or the weak statement and rated its strength on six items. Among the items were "I believe Urban-Furniture is concerned about my privacy" and "I believe that Urban-Furniture is concerned about the security of my financial information." Participants responded on a scale ranging from 1 ("disagree strongly") to 7 ("agree strongly"), and we averaged the responses ($\alpha = .94$). In support of the privacy/security manipulation, participants agreed more that UF would preserve their privacy and security when they read the strong privacy/security statement than when they read the weak statement (Ms = 5.56 versus 2.61; $F(1, 36) = 100.53$, $p < .01$).

Procedure. The procedure and survey were the same as in Study 1, with a few exceptions. To increase involvement, participants were told to imagine that they accepted a job in Manhattan after graduation and were searching for living-room furniture for their apartment. After viewing the site, participants in the strong and weak privacy/security conditions read these statements before completing the survey. To test whether reading such statements might artificially increase participants' risk perceptions of shopping online, we added seven items to the end of the survey that measured such concerns (see the Appendix). We averaged responses to provide a perceived risk score ($\alpha = .89$). Because the sample is homogeneous in terms of education and income, we deleted these demographic questions. In addition, because we measured online purchase experience in Study 1 with a single item that captures only recent online purchase experience, we added an item that measures general online purchase experience (i.e., how often participants shop online) on a scale from 1 ("not at all") to 7 ("quite often"). We averaged these items to capture online purchase experience ($r = .78$).

Results

Manipulation checks. In support of the investment manipulation, a 2 (investment) × 3 (privacy/security statement) ANOVA yielded a significant investment effect ($F(1, 73) = 6.84$, $p = .01$): Participants perceived the high-investment site as requiring a greater investment than the low-investment site (Ms = 3.36 versus 2.48). No other effects were significant ($Fs(1, 73) < 1$).

To examine whether reading a privacy/security statement might increase participants' perceived risks of shopping online, we compared risk perceptions using a 2 × 3 ANOVA. None of the effects were significant (the investment effect: $F(1, 73) = 1.22$, n.s.; the direct and interactive effects of the privacy/security statement: $Fs(2, 73) < 2.11$, n.s.). On average, participants perceived buying online as risky ($M = 5.71$, which is significantly higher than the midpoint of 4; higher numbers reflect greater perceived risk, $t(78) = 14.16$, $p < .01$).

Trusting beliefs and online purchase intentions. As in Study 1, we tested H_1 using hierarchical regression. As we predicted, ability beliefs influenced searchers' online purchase intentions ($\beta = .39$, $t(77) = 3.73$, $p < .005$; $R^2 = .15$, $F(1, 77) = 15.57$, $p < .01$, see Table 1). Furthermore, the addition of benevolence and integrity beliefs to the model did not significantly contribute to explaining online purchase intentions ($\Delta R^2 = .002$, $F(2, 75) < 1$). A stepwise regression analysis yielded the same results. Multicollinearity was not a problem (the maximum VIF = 2.58).

Web site investment and trusting beliefs. Here and elsewhere, we analyzed the data with a 2 (investment) × 3 (privacy/security statement) ANCOVA, controlling for online purchase experience and problems with the site. Ability beliefs were higher for those who visited the high-investment site than for those who visited the low-investment site (Ms = 2.68 versus 2.13; $F(1, 69) = 6.54$, $p < .05$), thus replicating the results of Study 1. No other effects on ability beliefs were significant ($Fs(1, 69) \leq 2.34$, n.s.). Furthermore, as we expected, Web site investment did not affect benevolence or integrity beliefs ($Fs(1, 68) < 2.29$, n.s.).

Whereas investment appears to communicate a firm's ability but not its benevolence and integrity, privacy/security statements appear to communicate benevolence and integrity but not ability: The privacy/security effect was significant for benevolence beliefs ($F(2, 69) = 3.30$, $p < .05$)

---

[2] Our analysis was limited to either strong or weak privacy and security statements. Thus, for ease of exposition, we use the term "privacy/security" to describe levels of privacy *and* security.

and integrity beliefs ($F(2, 69) = 4.03$, $p < .05$) but not for ability beliefs ($F(2, 69) = 2.34$, n.s.). Benevolence and integrity beliefs were significantly higher among those who read a strong statement than among those who read a weak statement (benevolence: Ms = 3.04 versus 2.65; $F(2, 68) = 2.89$, $p < .05$ [one-tailed test]; integrity: Ms = 3.19 versus 2.66; $F(2, 68) = 7.12$, $p < .05$). These beliefs were also higher for those who received a strong statement than for those who received no statement (benevolence: Ms = 3.04 versus 2.44; $F(2, 68) = 6.55$, $p < .05$; integrity: Ms = 3.19 versus 2.76; $F(2, 68) = 5.02$, $p < .05$). However, the mere presence of a privacy/security statement does not appear to signal a firm's benevolence and integrity; the difference between a weak statement and no statement was not significant at $p < .05$. It appears that both the presence and the strength of the statement signal the firm's benevolence and integrity.

Web site investment and online purchase intentions. We analyzed online purchase intentions with a 2 × 3 ANCOVA. In support of H_2, online purchase intentions were higher among those who visited the high-investment site than among those who visited the low-investment site (Ms = –.16 versus –.79; $F(1, 69) = 3.06$, $p < .05$ [one-tailed test]). A privacy/security statement did not moderate this effect ($F(2, 69) < 1$). Thus, regardless of whether benevolence and integrity signals (i.e., privacy/security statements) were present, ability signals (i.e., Web site investment) significantly influenced searchers' online purchase intentions.

The only other significant effect was a privacy/security effect ($F(1, 69) = 4.95$, $p = .01$). In support of the argument that benevolence and integrity signals (e.g., a strong privacy/security statement) should have little effect on searchers' online purchase intentions, online purchase intentions did not differ between those who received the strong privacy/security statement and those who received no statement (Ms = –.36 versus .09; $F(1, 46) = 1.10$, n.s.). However, online purchase intentions were lower among those who received a weak privacy/security statement than among those who received a strong statement (Ms = –1.15 versus –.36; $F(1, 41) = 3.82$, $p < .05$ [one-tailed test]) or no statement (Ms = –1.15 versus .09; $F(1, 49) = 8.58$, $p < .01$). This finding is consistent with existing research that negative cues regarding a firm's character tend to be unexpected, violating established norms about "business as usual" (Garfinkel 1963). In such situations, people appear unwilling to buy online from the firm.

To test whether ability beliefs mediate the investment effect on online purchase intentions (H_3), we added ability beliefs as a covariate to the 2 × 3 ANCOVA. Consistent with the requirements for mediation, ability was a significant covariate ($F(1, 68) = 8.47$, $p < .01$), and the investment effect became nonsignificant ($F(1, 68) < 1$). We replicated this finding with the Sobel (1982) test (Goodman I test statistic = 1.93; $p = .05$). As in Study 1, benevolence and integrity beliefs cannot be considered mediators, because investment did not significantly affect these beliefs. However, it is possible that benevolence and integrity beliefs mediate the privacy/security effect on online purchase intentions. To test this, we added these beliefs as covariates to the 2 × 3 ANCOVA. Inconsistent with the requirements for mediation but in support of the prediction that searchers' online purchase intentions are influenced by their ability rather than by their benevolence and integrity beliefs (H_1), neither of these beliefs were significant ($Fs(1, 67) < 2.02$, n.s.), and the privacy/security effect on online purchase intentions remained significant ($F(2, 67) = 5.75$, $p < .01$).

Conclusions

Replicating the results of Study 1, we found that searchers' ability beliefs, rather than their benevolence and integrity beliefs, influenced their online purchase intentions. As a result, ability signals (Web site investment) influenced their online purchase intentions more than did signals of the firm's benevolence and integrity (the presence of a strong privacy/security statement).

Because the objective of this research is to predict online purchase intentions, thus far our focus has been on people whose goals are most consistent with prepurchase deliberation, namely, searchers. For those with this goal, ability beliefs are a stronger driver of online purchase intentions than are benevolence and integrity beliefs. However, there may be goals that highlight the importance of a different component of trust than ability beliefs. Specifically, for those whose goal for visiting the site is more personal and less outcome oriented (i.e., browsers), a different pattern of results may emerge. Browsing is a moment-by-moment activity rather than a search process for a specific piece of information (Janiszewski 1998). As such, it is exploratory (Moe 2003) and reflects recreational behavior (Hoffman and Novak 1996). Similar to those who read a text for entertainment (Rosenblatt 1978), browsers are likely focused on what they are "living through" during their site visits. Thus, whereas searchers tend to be more objective and outcome oriented and thus are likely to disengage from having a personal experience with the site, browsers' experiences are likely more personal.

We argue that distinguishing between these goals has important implications for the relative impact of each trusting belief on online purchase intentions. Whereas searchers focus on performance and thus base their online purchase intentions on their ability beliefs, browsers likely have a more personal experience with the site and thus will be influenced by the most personal aspect of trust: benevolence beliefs. Benevolence beliefs reflect consumers' beliefs that the firm cares about their welfare and well-being (e.g., "I trust that the firm is concerned about my wants and needs, even if doing so results in profit reductions"), whereas integrity beliefs reflect beliefs about the firm's moral standards, regardless of how it feels about the individual (e.g., "I trust that the firm is guided by sound business principles and standards"). Likewise, ability beliefs reflect beliefs about the firm's expertise, regardless of how it feels about the individual (e.g., "I trust that the firm has the necessary skills to be successful"). Thus, because browsers' experiences are highly personal, their online purchase intentions should depend on the most personal dimension of trust (benevolence) rather than on the less personal dimensions (i.e., ability and integrity).

H_4: Browsers' online purchase intentions depend on their beliefs about the firm's benevolence rather than their beliefs about the firm's ability or integrity.

We also propose that the distinction between searching and browsing has important implications for the influence of Web site investment on online purchase intentions. Whereas prior research has demonstrated the importance of separating signals from interpretation, because not everyone interprets a signal in the same manner (Prabhu and Stewart 2001), it may be equally important to separate signal interpretation from response because even similarly interpreted signals may lead to meaningfully different responses. In particular, we argue that though both searchers and browsers will likely interpret investment as signaling the firm's abilities, the impact of this belief on their online purchase intentions will vary. Specifically, if searchers' online purchase intentions are affected by their ability beliefs, ability signals (or Web site investment) should influence their online purchase intentions. However, if browsers' online purchase intentions are unaffected by their ability beliefs, ability signals (even if interpreted as such) should have relatively little influence on their online purchase intentions. Consequently, we hypothesize the following:

H_5: Web site investment influences searchers' but not browsers' online purchase intentions.

Because the firm studied thus far was an unknown Internet-only firm, an additional objective of Study 3 was to replicate the results for searchers with a well-established firm that sells a different set of products (electronics rather than home furnishings and accessories).

Study 3

Method

Sample and design. A total of 152 undergraduate students participated in exchange for extra course credit. We randomly assigned participants to one of four conditions in a 2 (investment) × 2 (goal: searching versus browsing) experimental design.

Web site investment manipulation. In contrast to Studies 1 and 2, we manipulated investment using only technology. Specifically, for the high-investment site, we used Macromedia's Shockwave technology to allow participants to experience an online demonstration and to roll over the product image to gather additional information about its features. For the low-investment site, we conveyed the same information through text and static graphics rather than through the use of this technology.

Procedure. The procedure was similar to that used in Study 2, with a few exceptions. Each participant was seated at a computer terminal, which contained the instructions, the site, and the survey. Participants were told that they would be visiting a portion of Kodak's site devoted to a specific model of digital camera. Those assigned to browse were instructed to "have fun, looking at whatever you consider interesting and/or entertaining." Those assigned to search were instructed that before doing so, they should type two questions they have for Kodak about this digital camera. To ensure that both groups would be attending to information relevant to them, we did not specify what to look for. These instructions are identical to those used in prior research to instill a searching versus browsing goal (Schlosser 2003). Participants then visited the site. To control for the amount of time browsers versus searchers spent at the site and any possible impact of this on the dependent variables, all participants viewed the site for five minutes, which is comparable to the time imposed in prior research (Schlosser 2003). We did not investigate privacy/security statements in this study.

After visiting the site, participants completed the online survey, which was identical to that used in Study 2, except that the risk questions were replaced by the goal manipulation check. Participants were asked the extent to which their time at the site was spent looking for specific information (a searching activity) or looking to be entertained (a browsing activity) on a scale from 0 ("not at all") to 5 ("a lot"). In addition to the Web site investment items, participants rated how informative they found the site to be on a scale from 1 ("very little") to 7 ("a great deal"). We also made a slight change to the trusting beliefs measure: We asked participants to focus on the trustworthiness of Kodak's Internet marketing department. We made this change to direct participants' attention toward the e-commerce side of the firm rather than the firm in general or the brand. For example, participants may trust the offline firm's ability, benevolence, and integrity, but they may not believe that its Internet marketing managers share these same qualities.

Results

Manipulation checks. We analyzed the manipulation check items with a 2 (investment) × 2 (goal) ANOVA. In support of the goal manipulation, searchers reported spending more time searching (Ms = 2.52 versus 1.86; F(1, 150) = 14.08, $p < .01$) and less time browsing (Ms = 1.31 versus 1.97; F(1, 150) = 13.80, $p < .01$) than did browsers. In support of the investment manipulation, the high-investment site was perceived as a greater investment than the low-investment site (Ms = 4.51 versus 2.58; F(1, 148) = 88.76, $p < .01$). This effect emerged for both browsers (Ms = 4.15 versus 2.34; F(1, 74) = 20.19, $p < .01$) and searchers (Ms = 4.86 versus 2.10; F(1, 74) = 85.14, $p < .01$). No other effects were significant. The sites did not differ in perceived informativeness (F(1, 148) < 1).

Trusting beliefs and online purchase intentions. To test whether ability beliefs explain searchers' online purchase intentions (H_1) and whether benevolence beliefs explain browsers' online purchase intentions (H_4), we conducted separate hierarchical regression analyses for searchers and browsers; we modeled ability beliefs as a function of online purchase intentions (Model 1) before adding benevolence and integrity beliefs (Model 2). As we predicted, ability beliefs were significantly related to searchers' online purchase intentions ($\beta = .42$, t(74) = 3.96, $p < .01$; $R^2 = .18$, F(1, 74) = 15.70, $p < .01$; see Table 2). The addition of benevolence and integrity beliefs to the model did not significantly contribute to searchers' online purchase intentions ($\Delta R^2 = .03$, F(2, 72) = 1.50, n.s.). For browsers, ability

TABLE 2
Study 4: Hierarchical Regression Analysis of Trusting Beliefs on Online Purchase Intentions When Risk Is High Versus Low

A: Results from a Hierarchical Regression Analysis

		Model 1			Model 2		
Social Risk	**Variable**	β	R^2	F for R^2	β	ΔR^2	F for ΔR^2
High	Product attitudes	.60**	.53	7.91**	.58**	.02	1.80
	Autotelic NFT	–.09			–.14		
	Instrumental NFT	–.03			–.05		
	Online buying	.16			.06		
	Site problems	.10			.17		
	Ability	.29**			.24*		
	Integrity				.18		
Low	Product attitudes	.44**	.27	2.40**	.41**	.02	1.12
	Autotelic NFT	–.12			–.10		
	Instrumental NFT	.02			.00		
	Online buying	.05			.14		
	Site problems	.09			.01		
	Ability	.08			.17		
	Integrity				–.17		

B: Results from a Stepwise Regression Analysis

		Model 1			Model 2		
Social Risk	**Variable**	β	R^2	F for R^2	β	ΔR^2	F for ΔR^2
High	Product attitudes	.64**	.41	33.92**	.56**	.07	5.94**
	Ability				.27**		
Low	Product attitudes	.48**	.23	13.47**			

*$p < .10$.
**$p \leq .05$.
Notes: NFT = Need for touch.

beliefs alone did not significantly explain their online purchase intentions (R^2 = .02, F(1, 74) = 1.62, n.s.). However, the addition of benevolence and integrity beliefs to the model significantly contributed to browsers' online purchase intentions (ΔR^2 = .10, F(2, 72) = 4.28, $p < .05$). Consistent with H_4, browsers' benevolence beliefs were significantly related to their online purchase intentions (β = .33, t(74) = 2.70, $p < .01$), whereas ability and integrity beliefs were not (t(74) < 1). Multicollinearity was not a problem in the analyses (maximum VIFs < 1.50).

Web site investment, trusting beliefs, and online purchase intentions. We analyzed each belief with a 2 × 2 ANOVA. As we expected, ability beliefs were higher among those who visited the high-investment site than among those who visited the low-investment site (Ms = 3.51 versus 2.99; F(1, 148) = 21.74, $p < .01$). Investment did not affect benevolence and integrity beliefs (Fs(1, 148) < 1). Furthermore, people's goals for visiting the site neither directly affected nor interacted with investment to influence their ability, benevolence, or integrity beliefs (Fs(1, 148) < 2.27, n.s.). Thus, investment appears to signal a specific trusting belief (i.e., ability rather than benevolence or integrity), replicating the results of Studies 1 and 2 with a different set of stimuli. Moreover, this finding persisted despite differences in people's goals.

In support of H_5, searchers had higher online purchase intentions after visiting the high-investment site than after visiting the low-investment site (Ms = –.03 versus –.76; F(1, 74) = 4.92, $p < .05$), whereas browsers' online purchase intentions were unaffected by investment (Ms = –.48 versus –.08 at the high- and low-investment sites; F(1, 74) = 1.26, n.s.). This investment × goal interaction was significant (F(1, 148) = 5.56, $p < .05$).

In addition to demonstrating that the site effect on online purchase intentions varies across goals, we tested whether the mediating effects of ability vary across goals using the procedure that Baron and Kenny (1986) outline. Specifically, we regressed online purchase intentions on goal, site, goal × ability, and site × ability, and we found a significant goal × ability effect (t(146) = 1.82, $p < .05$ [one-tailed test]). For searchers, mediation was supported: Ability was significant (F(1, 73) = 10.45, $p <$ 01), and the investment effect became nonsignificant (F(1, 73) < 1). A Sobel (1982) test supports this finding (Goodman I test statistic = 2.74, $p < .01$). However, for browsers, mediation was not supported: Ability had little effect on online purchase intentions (F(1, 73) = 2.64, n.s.). Furthermore, because investment did not affect benevolence and integrity beliefs, they cannot be considered mediators.

Conclusions

In Study 3, we examined a boundary condition for the effect of Web site investment on online purchase intentions: consumers' goals for visiting sites. We found that the effects of Web site investment and ability on online purchase intentions are specific to searchers and do not generalize to browsers. For browsers, the most personal component of trust (i.e., benevolence rather than ability) influences their online purchase intentions. Consequently, although browsers recognized that Web site investment signals ability, it had relatively little influence on their online purchase intentions.

A remaining question is whether the findings are really a matter of trust. Recall that trusting intentions involve risk (Moorman, Zaltman, and Deshpandé 1992), which causes people to consult their trusting beliefs to determine whether to perform the trusting behavior. If the findings are driven by trust, Web site investment should influence consumers' online purchase intentions under conditions of risk. Although buying online is typically risky, certain situational factors (e.g., buying a gift for a significant versus an insignificant other) can make buying online more or less risky. When there is relatively little risk, buying online should involve relatively little trust. Consequently, people are less likely to consult their trusting beliefs when deciding how to act. Indeed, the degree of trust necessary to influence behavior has an approximate linear relationship to the degree of risk involved (Corritore, Kracher, and Wiedenbeck 2003; Mayer, Davis, and Schoorman 1995). Thus, when the purchasing scenario involves relatively little risk, ability beliefs (and signals designed to influence such beliefs) should have relatively little effect on searchers' online purchase intentions.

H_6: Ability beliefs affect searchers' online purchase intentions only when buying involves risk.

H_7: Web site investment affects searchers' online purchase intentions when buying involves risk.

Another objective for Study 4 was to test the importance of trust in determining online purchase intentions versus online purchase experience and a desire to examine products physically before purchase. For people who are motivated to touch products, barriers to touch can decrease confidence in product evaluation, though conveying haptic information through text and graphics can help (Peck and Childers 2003b). If trust is critical in determining online purchase intentions, ability beliefs should be related to online purchase intentions even when we account for these other factors.

Study 4

Method

Sample and design. A total of 98 undergraduate students participated in exchange for extra course credit. We randomly assigned them to one of four conditions in a 2 (investment: high- versus low-investment site) × 2 (risk: high versus low) factorial design. They visited the UF site used in Studies 1 and 2.

Risk manipulation. We manipulated risk by varying social risk. Specifically, the high-risk scenario was as follows: "Imagine that you graduated and obtained your dream job at a firm. You are invited to your boss' house-warming party, which takes place in a week." The low-risk scenario was as follows: "Imagine that you are invited to a former roommate's house-warming party, which takes place in a month." Participants in both scenarios were told that this person likes modern furniture and accessories, such as those offered at UF's site, and that they intend to buy a home accessory (e.g., a vase) as a house-warming gift for this person.

To test the effectiveness of these scenarios, 41 undergraduate students read either scenario and then rated how nervous and concerned they would be about making this purchase and how risky they considered this purchase on a scale from 1 to 7 (higher numbers reflected greater risk). We averaged these items to form a perceived risk score ($\alpha = .76$). As we expected, perceived risk was higher among those who read the high-risk scenario than among those who read the low-risk scenario (Ms = 4.22 versus 3.12; $F(1, 37) = 7.12$, $p = .01$). Participants then looked at screen captures of the high- or low-investment site and answered the Web site investment items ($\alpha = .94$; see the Appendix). As we expected, perceived investment was higher for the high-investment site than for the low-investment site (Ms = 5.12 versus 3.81; $F(1, 37) = 14.78$, $p < .01$). Moreover, risk did not directly affect or moderate these investment perceptions ($Fs(1, 37) < 1.39$, n.s.).

Procedure. The procedure and survey were the same as that used in Study 2, with a few exceptions. At the beginning of the experiment, participants read either the high- or the low-risk scenario before receiving the search instructions. Participants did not receive a privacy/security statement.

Unlike Study 2, online purchase intentions were specific to home accessories (the gift they were to buy). Participants also reported their attitudes toward UF's home accessories to control for any individual differences in their liking of UF's offerings (see the Appendix). In addition, to test the robustness of our effects related to ability versus integrity beliefs, we replaced Mayer and Davis's (1995) scale with an adaptation of Moorman, Deshpandé, and Zaltman's (1993) ability and integrity items.[3] Three items measured ability beliefs ($\alpha = .93$), and two items measured integrity beliefs ($r = .31$; see the Appendix). Finally, to test the impact of trust on online purchase intentions relative to consumers' desire to touch products before purchase, we administered Peck and Childers's (2003a) need for touch (NFT) scale, which measures the need for autotelic and instrumental touch (see the Appendix).

Results

Trusting beliefs and online purchase intentions. To test H_6, we conducted a hierarchical regression analysis separately for searchers in the high- and low-risk conditions; we modeled ability beliefs and variables believed to influence online purchase intentions (attitudes toward home accessories, both types of NFT, online purchase experience, and site problems) as a function of online purchase intentions (Model 1) before adding integrity beliefs (Model 2). As we predicted, for searchers in the high-risk condition, ability beliefs were significantly related to online purchase intentions ($\beta = .29$, $t(48) = 2.39$, $p < .05$; $R^2 = .52$, $F(6, 43) = 7.91$, $p < .01$; see Table 2). The only other significant variable was attitudes toward home accessories ($\beta = .60$, $t(48) =$

---

[3] Moorman, Deshpandé, and Zaltman (1993) do not measure benevolence beliefs; thus, we do not address them here.

5.06, $p < .01$). The addition of integrity beliefs to the model did not significantly contribute to explaining online purchase intentions ($\Delta R^2 = .02$, $F(1, 42) < 1$). For searchers in the low-risk condition, the only significant variable from Model 1 was attitudes toward home accessories ($\beta = .44$, $t(45) = 2.88, p < .01$; $R^2 = .27$, $F(6, 40) = 2.40, p < .05$). The addition of integrity beliefs did not significantly contribute to explaining online purchase intentions ($\Delta R^2 = .02$, $F(1, 39) = 1.12$, n.s.). Consistent with H_6, ability beliefs influenced online purchase intentions only when risk was high. When risk was low, neither ability nor integrity beliefs could explain online purchase intentions. Moreover, ability beliefs is an important variable in explaining searchers' trusting intentions: In the high-risk condition, ability beliefs explained more variance in their intentions to buy online than did individual differences in NFT or online purchase experience. A stepwise regression led to the same conclusions. Furthermore, multicollinearity was not an issue: The maximum VIFs were below 1.99.

Web site investment, trusting beliefs, and online purchase intentions. For both beliefs, we conducted a 2 × 2 ANCOVA. As we expected, ability beliefs were higher among those who visited the high-investment site than among those who visited the low-investment site (Ms = 4.01 versus 2.84; $F(1, 90) = 14.82, p < .01$). Investment did not affect integrity beliefs ($F(1, 90) = 3.30$, n.s.). Furthermore, risk had neither a direct nor a moderating effect on either belief ($Fs(1, 90) < 1$). Thus, regardless of purchase risk, investment communicates a firm's ability.

For online purchase intentions for home accessories (the gift), a 2 × 2 ANCOVA yielded a significant investment × risk interaction ($F(1, 88) = 6.06, p < .05$). Consistent with H_7, when risk was high, searchers' online purchase intentions were higher at the high-investment site than at the low-investment site (Ms = .76 versus –.38; $t(49) = 2.64, p < .05$), whereas when risk was low, investment had little effect on their online purchase intentions (Ms = .46 versus .82 at the high- and low-investment sites; $t(47) < 1$).

In addition to demonstrating that the site effect on online purchase intentions varies across risk, we tested whether the mediating effects of ability vary across levels of risk by regressing online purchase intentions on risk, site, risk × ability, and site × ability. As we expected, the risk × ability effect was significant ($t(96) = 1.93, p = .057$). In the high-risk condition, mediation was supported: Ability was a significant covariate ($F(1, 43) = 8.67, p < 01$), whereas the previously significant site effect was reduced to nonsignificance ($F(1, 43) < 1$). The Sobel (1982) test provides further support for this finding (Goodman I test statistic = 2.13, $p < .05$). In the low-risk condition, however, mediation was not supported: It did not affect online purchase intentions ($F(1, 40) < 1$). As in Studies 1–3, integrity cannot be considered a mediator, because investment did not significantly affect this variable.

Discussion

Collectively, these studies clarify the role of trust in predicting online purchase intentions. Across four studies with different samples, companies, and products sold, we demonstrate that Web site investment influences searchers' intentions to buy online by influencing one component of trusting beliefs, ability (versus benevolence and integrity). Five conclusions can be drawn from our findings: First, they provide empirical support for the position that trusting beliefs should be considered separately from trusting intentions (Moorman, Deshpandé, and Zaltman 1993; Moorman, Zaltman, and Deshpandé 1992). Although definitions of trust typically contain both (e.g., Rotter 1980), most empirical research on consumer trust implicitly assumes that measuring trusting beliefs alone is sufficient (e.g., Morgan and Hunt 1994). However, our research suggests that it is important to measure trusting intentions because building trusting beliefs does not necessarily lead to higher trusting intentions. For example, although the presence (versus absence) of a strong privacy/security statement increased searchers' benevolence and integrity beliefs in Study 2, it did not lead to higher online purchase intentions. Likewise, in Study 3, Web site investment increased browsers' ability beliefs but not their online purchase intentions.

Second, our findings are consistent with the view that risk is a necessary (but not a sufficient) condition for trust (Moorman, Deshpandé, and Zaltman 1993; Moorman, Zaltman, and Deshpandé 1992). Specifically, in Study 4, we demonstrate that Web site investment and trusting beliefs influence online purchase intentions only when risk is high, or when the behavior requires trust. In contrast, when risk is low (and, thus, the behavior does not require trust), online purchase intentions were largely unaffected by Web site investment and trusting beliefs.

Third, we demonstrate the importance of considering the differential impact of trusting beliefs on trusting intentions. As we demonstrated across four studies, trusting beliefs do not have an equally strong influence on online purchase intentions. For example, in Study 2, we find that among searchers, gains in ability beliefs led to higher online purchase intentions, whereas gains in benevolence and integrity beliefs did not. Consequently, signals that were effective in building benevolence and integrity beliefs (i.e., the presence of a strong privacy/security statement) were less effective than ability signals (i.e., Web site investment) in increasing online purchase intentions. Together, these findings suggest that it is important for an online firm to identify the most influential trusting belief in order to identify the signal that will most effectively increase trusting intentions.

Fourth, our research contributes to the signaling literature. Prior research has demonstrated the importance of separating signals from interpretation because not all consumers interpret signals in the same way (Prabhu and Stewart 2001). Our research builds on this by demonstrating that even when consumers interpret signals as a firm intends, their responses may vary significantly. For example, although searchers and browsers (Study 3) and those perceiving high and low risk (Study 4) interpreted Web site investment as a signal of ability, Web site investment significantly affected only the online purchase intentions of searchers and for those perceiving high risk.

Fifth, these findings are also helpful in explaining what appear to be contradictory findings regarding the effectiveness of privacy/security statements relative to Web site design. Our research suggests that though privacy/security statements can influence certain trusting beliefs (i.e., benevolence and integrity beliefs), they are less effective in increasing searchers' willingness to buy online. Consequently, even when privacy/security statements improve beliefs about a firm's trustworthiness, they can still be relatively less effective in converting visitors to buyers. These findings also help explain why Web site design plays such an important role in online purchase intentions. Instead of serving a purely aesthetic function, it signals that a firm's ability can be trusted, which we found to be the most significant driver of searchers' online purchase intentions.

Managerial Implications

The prevailing wisdom has been that concerns about buying online would diminish as consumers gain experience buying online (eMarketer 2005). Thus, marketing managers may wonder whether they need to invest in establishing trust as consumers gain online purchase experience. Contrary to this prevailing wisdom, however, recent reports indicate that consumers at all levels of Internet experience are increasingly cautious about buying online (eMarketer 2005; Forrester 2005). Compounding this problem is the notion that the actual risks associated with being online may even increase over time (Rust, Kannan, and Peng 2002). Consistent with the view that firms need to establish trust regardless of their visitors' online buying experience at other sites, we found that Web site investment had a significant effect on consumers' online purchase intentions by influencing their ability beliefs even when we controlled for online purchase experience (Studies 1–4).

Likewise, marketing managers of established offline firms may wonder whether they need to invest in Web site design to establish trust online. Our findings suggest that establishing trust is important for both an established offline firm (Kodak) and an unknown pure play (UF). It may seem surprising that consumers did not trust the ability of a known firm to handle online transactions successfully, yet this finding may reflect well-founded consumer skepticism about whether firms have the same abilities online as they do offline. Indeed, recently publicized examples of delivery failures by known firms (e.g., Toys "R" Us during the 2003 Christmas season) likely contribute to such concerns. Part of consumers' caution may also stem from their awareness that offline firms may outsource e-commerce functions, thus making it difficult to predict the quality of the online transaction on the basis of their offline experiences with the firm.

To use Web site investment effectively, however, managers should identify the reason most consumers visit their site. Our results indicate that Web site investment is effective in boosting online purchase intentions when visitors are searchers. Thus, if most visitors are searchers, Web site investment is warranted, but if most visitors are browsers, such investments may be ineffective. If the site attracts both, it may be best to invest mainly in pages that are most likely to attract searchers. For example, compared with browsers, searchers tend to "drill down" beyond category-level pages (e.g., a page featuring all the living-room furniture that a store offers) to specific product pages (e.g., the pages on the Bay Club chair; Moe 2003). Thus, investment in such technology as Shockwave or Design Within Reach may be most effective at the product level rather than the category (or home page) level. Firms can identify the segment visiting their sites by using clickstream data (searchers spend more time at product and search pages than browsers do; Moe 2003), usability studies, or more traditional market research techniques, such as online surveys or panel studies.

It appears that Web site investment is effective for purchases that involve risk, particularly social risk. For example, our findings suggest that sites selling gifts for important social occasions (e.g., weddings, birthdays, formal business functions) or holidays (e.g., Christmas, Valentine's Day) may benefit more from Web site investment than sites selling products with less risk (e.g., office supplies). For firms selling products that vary in social risk, investing in Web site design would likely be most effective at the product-level pages for products that are most likely to carry high social risk (e.g., jewelry).

Our focus has been on how to convert visitors into those who are willing to buy online. However, the firm may have other objectives, such as creating goodwill. Because Web site investment did not significantly affect benevolence or integrity beliefs in any of our studies, this tactic is likely to be ineffective in meeting such objectives. The findings from Study 3 suggest that privacy/security statements are better than Web site investment for influencing these trusting beliefs. It would be valuable to examine how other signals (e.g., announcements of investments in social causes) might communicate benevolence and integrity as well.

This research also has important public policy implications. Although signals can be used to communicate a firm's true abilities and intentions, they can also be used to mislead (Prabhu and Stewart 2001). Unscrupulous retailers may take advantage of consumers by offering high-investment sites for inferior products or an inferior online transaction or, worse, by stealing the consumer's identity. Indeed, according to the National Consumer League (2005), the second most common Internet scam complaint was about products that were never delivered or were misrepresented at an e-commerce site. Yet many of the Federal Trade Commission's (FTC's) initiatives focus on e-commerce problems at the technological level by recommending better encryption and firewall technology (e.g., FTC 2003). Although such investments are important, our findings suggest that e-commerce problems are a marketing issue as well. Specifically, our research indicates that consumers infer from investments in site design that the firm's abilities can be trusted. As a result, they are more willing to buy from the firm. Although a firm with a well-designed site has demonstrated an investment in site design, it does not follow that the firm itself has these capabilities. Oftentimes, site design and order fulfillment are outsourced and thus do not reflect the firm's online abilities. Furthermore, even if the firm developed the site internally, it does not follow that its abilities in this area will generalize to other areas, such as order fulfillment. However, because consumers appear to infer that a firm with a well-designed site can be trusted, unscrupulous firms could use Web site investment to deceive through implication.

As with regulating other e-commerce practices, there is debate about whether government intervention is necessary or whether industry self-regulation is sufficient (Rust, Kannan, and Peng 2002). Currently, the FTC has taken a caveat emptor stance toward such issues. Although the FTC provides standards for what to include in a privacy/security statement (e.g., notifying consumers of a firm's information practices and ensuring the security of the information collected) and advises consumers to consult these statements before buying (www.ftc.gov), our results suggest that even when consumers read these statements, their online purchase intentions are largely unaffected by them unless they are noticeably weak. Instead, consumers' intentions to buy online from a vendor are affected more by its Web site investment. Thus, the FTC's suggestions may not fully protect consumers from being victimized. Conversely, industry self-regulation may be sufficient. Investing in Web site design may not be profitable for most low-quality firms, thus discouraging such firms from doing so. Publicly visible signals that require high up-front investments before any sales transactions rely on repeat purchases and word-of-mouth recommendations to recoup such sunk costs (Kirmani and Rao 2000). Thus, even if low-quality firms can induce trial with a high-investment site, the firm's true abilities would be revealed. Consequently, future purchases would be unlikely, making it difficult for low-quality firms to recover Web site expenditures.

Limitations and Future Directions

As with any study, the findings should be considered in light of their limitations. For example, instead of being part of the site, the privacy/security statements used were included in the survey booklets. This enabled us (1) to ensure that those in the strong and weak privacy/security conditions were exposed to the statement and (2) to control the timing of when they read the statement relative to experiencing the site. Although such gains in control come at the expense of face validity, this likely provided a more conservative test of our predictions. That is, by placing the statements in the survey booklet and exposing participants to them after their visit to the site (but immediately before reporting their online purchase intentions), we likely increased their salience. Yet despite this, investment had a stronger effect on online purchase intentions than did the presence of a strong privacy/security statement. Still, it would be beneficial to examine what, if any, effect a privacy/security statement would have if its exposure was allowed to vary. Perhaps the first hurdle is to establish trust in the firm's ability. Only then might searchers consult such statements.

Although we studied several products (i.e., furniture and accessories, which are likely to be evaluated more on their experiential qualities than on their search qualities, and digital cameras, which are likely to be evaluated more on their search qualities than on their experiential qualities), all were high-ticket items. It would be fruitful to examine whether such results generalize to inexpensive items, such as CDs or books. For such products, the risk of buying online might be reduced, thus reducing the effect of investment on online purchase intentions. It would also be fruitful to examine whether Web site investment has the greatest influence on trusting intentions for sites that feature services, such as banking or travel. When transactions are more personal in nature, perhaps trust in the firm's benevolence would be as important as, if not more important than, trust in its ability.

Note that participants did not have prior experience buying online from either UF or Kodak (which did not sell directly from its site at the time of Study 3). It would be worthwhile to examine how online purchase experience with a firm might influence consumers' interpretations of the firm's signals, as well as their trusting beliefs and intentions. Furthermore, we examined only signals that can be observed and experienced, but other types of signals are also possible (Kirmani and Wright 1989). For example, firms might signal their Web site investments through announcements in press releases or on the site itself. However, because such announcements require relatively little cost (and thus could be used by both high- and low-quality firms), they may be less effective than site investment in building trust. This would be a worthwhile future research direction.

Conclusion

Despite these limitations, we believe that our framework provides meaningful insights into how firms can most effectively signal their trustworthiness to convert online visitors to buyers. Furthermore, it contributes to the trust literature by demonstrating the importance of treating trusting beliefs as (1) a multidimensional construct and (2) a necessary but not sufficient condition for trust. Similarly, it contributes to the signaling literature by demonstrating that even if consumers interpret signals as a firm intends, their responses can vary significantly.

APPENDIX
Measures and Items

Measures and Items	Source
Online Purchase Intentions (Trusting Intentions)	New scale
"Unlikely/likely" (–3 to +3)	
"Impossible/possible" (–3 to +3)	
"Improbable/probable" (–3 to +3)	
Trusting Beliefs	
Ability (scale used in Studies 1–3)[a, b]	Mayer and Davis (1999)
Urban-Furniture seems very capable of performing online transactions.	
Urban-Furniture appears to be successful at the things it tries to do.	
Urban-Furniture seems to have much knowledge about what needs to be done to fulfill online transactions.	
I feel very confident about Urban-Furniture's online skills.	
Urban-Furniture appears to have specialized capabilities that can increase its performance with online transactions.	
Urban-Furniture appears to be well qualified in the area of e-commerce.	
Benevolence (scale used in Studies 1–3)[a, b]	Mayer and Davis (1999)
Urban-Furniture seems very concerned about my welfare.	
My needs and desires appear to be important to Urban-Furniture.	
It doesn't seem that Urban-Furniture would knowingly do anything to hurt me.	
Urban-Furniture seems to really look out for what is important to me.	
Urban-Furniture appears to go out of its way to help me.	
Integrity (scale used in Studies 1–3)[a, b]	Mayer and Davis (1999)
Urban-Furniture seems to have a strong sense of justice.	
Urban-Furniture appears to try hard to be fair in dealings with others.	
Urban-Furniture's actions and behaviors are not very consistent. (reverse scored)	
I like Urban-Furniture's values.	
Sound principles seem to guide Urban-Furniture's behavior.	
Ability (scale used in Study 4)	Moorman, Deshpandé, and Zaltman (1993)
"Nonexpert/expert" (1 to 7)	
"Untrained/trained" (1 to 7)	
"Inexperienced/experienced" (1 to 7)	
Integrity (scale used in Study 4)	Moorman, Deshpandé, and Zaltman (1993)
"No integrity/integrity" (1 to 7)	
Urban-Furniture does not have a great deal of integrity.[c] (reverse scored)	
Web Site Investment	New scale
The amount of time invested into developing this website seems to be: "very little/a great deal." (1 to 7)	
The amount of effort devoted to developing this website seems to be: "very little/a great deal." (1 to 7)	
The amount of money invested into developing this website seems to be: "very little/a great deal." (1 to 7)	
Online Risk Perceptions *(scale used in Study 2)*[c]	New scale
Shopping online is risky.	
Providing credit card information online is risky.	
Providing personal information (i.e., social security number and mother's maiden name) online is risky.	
Purchasing items online is risky.	
Providing my e-mail address and phone number online is risky.	
Registering online is risky.	
It is riskier to shop online for a product than to shop offline for it.	
Attitudes Toward Home Accessories *(scale used in Study 4)*	New scale
"Bad/good" (–3 to +3)	
"Unpleasant/pleasant" (–3 to +3)	
"Dislike/like" (–3 to +3)	

APPENDIX
Continued

Measures and Items	Source
NFT *(scale used in Study 4)*[c]	Peck and Childers (2003a)
Autotelic NFT	
Touching products can be fun.	
I like to touch products even if I have no intention of buying them.	
When browsing in stores, I like to touch a lot of products.	
When walking through stores, I can't help touching all kinds of products	
When browsing in stores, it is important for me to handle all kinds of products	
I find myself touching all kinds of products in stores.	
Instrumental NFT	
I place more trust in products that can be touched before purchase.	
I feel more comfortable purchasing a product after physically examining it.	
I feel more confident making a purchase after touching a product.	
If I can't touch a product in the store, I am reluctant to purchase the product.	
The only way to make sure a product is worth buying is to actually touch it.	
There are many products that I would only buy if I could handle them before purchase.	

[a] Measured on a scale from 1 ("disagree strongly") to 5 ("agree strongly").
[b] In Study 3, we replaced Urban-Furniture with Kodak.
[c] Measured on a scale from 1 ("disagree strongly") to 7 ("agree strongly").

REFERENCES

Bargh, John A. and Katelyn Y.A. McKenna (2004), "The Internet and Social Life," *Annual Review of Psychology*, 55 (February), 573–90.

Baron, Reuban M. and David A. Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (December), 1173–82.

Corritore, Cynthia L., Beverly Kracher, and Susan Wiedenbeck (2003), "Online Trust: Concepts, Evolving Themes, A Model," *International Journal of Human-Computer Studies*, 58 (June), 737–58.

Doney, Patricia M. and Joseph P. Cannon (1997), "An Examination of the Nature of Trust in Buyer–Seller Relationships," *Journal of Marketing*, 61 (April), 35–51.

Dunn, Julie (2004), "Survey Shows Online Security Perception Gap Between Experts, Users," *Knight Ridder Tribune Business News*, (August 17), 1.

eMarketer (2005), "Privacy and Security: Fraud, Identity Theft, Phishing, Viruses and Other Threats," (accessed April 2, 2005), [available at http://www.emarketer.com/Report.aspx?privacy_mar05].

Fogg, B.J., Cathy Soohoo, David Danielson, Leslie Marable, Julianne Stanford, and Ellen Tauber (2002), How Do Users Evaluate the Credibility of Web Sites? A Study with over 2,500 Participants," (accessed December 6, 2005), [available at http://portal.acm.org/citation.cfm?doid=997078.997097?].

Forrester, Jonathan (2005), "Rebuilding Consumers' Trust in the Internet," (accessed April 2, 2005), [available at http://www.forrester.com/Research/Document/Excerpt/0,7211,35612,00.html].

FTC (2003), "FTC Working to Protect Consumers and Businesses from Information Security Breaches," press release, (November 19), (accessed December 6, 2005), [available at http://www.ftc.gov/opa/2003/11/cybersecurity.htm].

Ganeson, Shankar (1994), "Determinants of Long-Term Orientation in Buyer–Seller Relationships," *Journal of Marketing*, 58 (April), 1–19.

Garfinkel, Harold (1963), "A Conception of, and Experiments with, 'Trust' as a Condition of Stable Concerted Actions," in *Motivation and Social Interaction*, O.J. Harvey, ed. New York: Ronald Press, 187–239.

Hoffman, Donna L. and Thomas P. Novak (1996), "Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations," *Journal of Marketing*, 60 (July), 50–68.

——— and Marcos A. Peralta (1999), "Information Privacy in the Marketspace: Implications for the Commercial Uses of Anonymity on the Web," *The Information Society*, 15 (2), 129–39.

Janiszewski, Chris (1998), "The Influence of Display Characteristics on Visual Exploratory Searching Behavior," *Journal of Consumer Research*, 25 (December), 290–312.

Kim, Peter H., Donald L. Ferrin, Cecily D. Cooper, and Kurt T. Dirks (2004), "Removing the Shadow of Suspicion: The Effects of Apology Versus Denial for Repairing Competence-Versus Integrity-Based Trust Violations," *Journal of Applied Psychology*, 89 (1), 104–118.

Kirmani, Amna and Akshay R. Rao (2000), "No Pain, No Gain: A Critical Review of the Literature on Signaling Unobservable Product Quality," *Journal of Marketing*, 64 (April), 66–80.

——— and Peter Wright (1989), "Money Talks: Perceived Advertising Expense and Expected Product Quality," *Journal of Consumer Research*, 16 (December), 344–53.

Kumar, Nirmalya, Lisa K. Scheer, and Jan-Benedict E.M. Steenkamp (1995), "The Effects of Supplier Fairness on Vulnerable Resellers," *Journal of Marketing Research*, 32 (February), 54–65.

Mayer, Roger and James Davis (1999), "The Effect of the Performance Appraisal System on Trust for Management: A Field Quasi-Experiment," *Journal of Applied Psychology*, 84 (February), 123–36.

———, ———, and F. David Schoorman (1995), "An Integrative Model of Organizational Trust," *Academy of Management Review*, 20 (July), 709–734.

McKnight, D. Harrison, Vivek Choudhury, and Charles Kacmar (2002), "Developing and Validating Trust Measures for e-Commerce: An Integrative Typology," *Information Systems Research*, 3 (September), 334–59.

Moe, Wendy W. (2003), "Buying, Searching, or Browsing: Differentiating Between Online Shoppers Using In-Store Navigational Clickstream," *Journal of Consumer Psychology*, 13 (1–2), 29–39.

Montoya-Weiss, Mitzi M., Glenn B. Voss, and Dhruv Grewal (2003), "Determinants of Online Channel Use and Overall Satisfaction with a Relational, Multichannel Service Provider," *Journal of the Academy of Marketing Science*, 31 (Fall), 448–58.

Moorman, Christine, Rohit Deshpandé, and Gerald Zaltman (1993), "Factors Affecting Trust in Market Research Relationships," *Journal of Marketing*, 57 (January), 81–101.

———, Gerald Zaltman, and Rohit Deshpandé (1992), "Relationships Between Providers and Users of Market Research: The Dynamics of Trust Within and Between Organizations," *Journal of Marketing Research*, 29 (August), 314–28.

Morgan, Robert M. and Shelby D. Hunt (1994), "The Commitment–Trust Theory of Relationship Marketing," *Journal of Marketing*, 58 (July), 20–39.

Naquin, Charles E. and Gaylen D. Paulson (2003), "Online Bargaining and Interpersonal Trust," *Journal of Applied Psychology*, 88 (1), 113–20.

National Consumer League (2005), *Internet Fraud Statistics: Internet Scams. Fraud Trends: January–June 2005*, (accessed December 6, 2005), [available at http://www.fraud.org/internet/internet_scams_halfyear_2005.pdf].

Neter, John, William Wasserman, and Michael H. Kutner (1990), *Applied Linear Statistical Models*. Boston: Richard D. Irwin.

Palmer, Jonathan W., Joseph P. Bailey, and Samer Faraj (2000), "The Role of Intermediaries in the Development of Trust on the WWW: The Use and Prominence of Trusted Third Parties and Privacy Statements," *Journal of Computer-Mediated Communication*, 5 (March), 1–25.

Peck, Joann and Terry L. Childers (2003a), "Individual Differences in Haptic Information Processing: The 'Need for Touch' Scale," *Journal of Consumer Research*, 30 (December), 430–42.

——— and ——— (2003b), "To Have and To Hold: The Influence of Haptic Information on Product Judgments," *Journal of Marketing*, 67 (April), 35–48.

Porter, Michael E. (1980), *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. New York: The Free Press.

Prabhu, Jaideep and David W. Stewart (2001), "Signaling Strategies in Competitive Interaction: Building Reputations and Hiding the Truth," *Journal of Marketing Research*, 38 (February), 62–73.

Rosenblatt, Louise (1978), *The Reader, the Text, the Poem: The Transactional Theory of the Literary Work*. Carbondale: Southern Illinois University Press.

Rotter, Julian B. (1980), "Interpersonal Trust, Trustworthiness, and Gullibility," *American Psychologist*, 35 (January), 1–7.

Rust, Roland, P.K. Kannan, and Na Peng (2002), "The Customer Economics of Internet Privacy," *Journal of the Academy of Marketing Science*, 30 (4), 455–64.

Schlosser, Ann E. (2003), "Experiencing Products in a Virtual World: The Role of Goals and Imagery in Influencing Attitudes Versus Intentions," *Journal of Consumer Research*, 30 (3), 184–98.

Sirdeshmukh, Deepak, Jagdip Singh, and Barry Sabol (2002), "Consumer Trust, Value, and Loyalty in Relational Exchanges," *Journal of Marketing*, 66 (January), 15–27.

Sitkin, Sim B. and Nancy L. Roth (1993), "Explaining the Limited Effectiveness of Legalistic 'Remedies' for Trust/Distrust," *Organization Science*, 4 (August), 367–92.

Smith, J. Brock and Donald W. Barclay (1997), "The Effects of Organizational Differences and Trust on the Effectiveness of Selling Partner Relationships," *Journal of Marketing*, 61 (January), 3–21.

Sobel, M.E. (1982), "Asymptotic Intervals for Indirect Effects in Structural Equation Models," in *Sociological Methodology*, S. Leinhart, ed. San Francisco: Jossey-Bass, 290–312.

Sultan, Fareena, Glen L. Urban, Venkatesh Shankar, and Iakov Y. Bart (2002), "Determinants and Role of Trust in E-Business: A Large Scale Empirical Study," Working Paper No. 4282, Sloan School of Management, Massachusetts Institute of Technology.

Weiner, Bernard (1972), *Theories of Motivation: From Mechanism to Cognition*. Chicago: Rand McNally.

——— (1986), *An Attributional Theory of Motivation and Emotion*. New York: Springer-Verlag.

ANNOUNCEMENT

CALL FOR NOMINATIONS

2006 JAGDISH N. SHETH LEADERSHIP MEDAL

The American Marketing Association and the Sheth Foundation are pleased to solicit nominations for the **2006 Jagdish N. Sheth Leadership Medal**. The Sheth Leadership Medal will be awarded annually to a current or former Chief Executive Officer (CEO) who has made long-term, transformational contributions to the practice and discipline of marketing through his or her vision, innovativeness, and leadership in creating value for customers. It will be presented at the American Marketing Assocation’s new industry meeting, Mplanet, in November 2006.

Nominees must possess a reputation that reflects their passion for customers and must have become a role model for future generations of industry leaders.

NOMINATIONS SHOULD BE SUBMITTED TO:

Sheth Leadership Medal
Blue Ribbon Panel
William H. Cunningham, Chair
P.O. Box E
The University of Texas at Austin
Austin, TX 78713
whc@po.utexas.edu
512.232.7540

Additional information and nomination forms can be obtained from the chair of the panel. Nominations will be reviewed and the recipient selected by the Sheth Leadership Medal Blue Ribbon Panel. ***Nominations must be submitted no later than April 30, 2006.***

AMERICAN
MARKETING
ASSOCIATION

You have to see it to believe it.

It's a marketing event like no other.

A journey of **unparalleled discovery** offering ground-breaking solutions to today's most pressing marketing challenges.

And you're invited.

Join the **American Marketing Association** and step outside the realm of the typical marketing conference:

> **EXPLORE** the new frontier of marketing for fresh insights and never-before-seen research on connecting with customers.

> **TRANSFORM** your marketing organization and change mindsets with innovative tools and techniques that drive actionable results enterprise wide.

> **CONNECT** with peers in a podium-free environment focused on collaboration and interaction.

> **EXPERIENCE** over two days of total marketing immersion with industry leaders and rising stars on an unprecedented journey of knowledge and inspiration.

M planet™
AMERICAN MARKETING ASSOCIATION

Start your journey today: **Mplanet2006.com**

November 29–December 1, 2006
Walt Disney World Dolphin Resort

Journal of Marketing

July 2006
Volume 70, Number 3

A Quarterly Publication of the
American Marketing Association

From the Editor: The Maturation of Marketing as an Academic Discipline 1 Roland T. Rust

From the Web Editor: Improvements to the *Journal of Marketing* Web Site 3 Siva K. Balasubramanian

Valuing the Real Option of Abandoning Unprofitable Customers When Calculating Customer Lifetime Value 5 Michael Haenlein, Andreas M. Kaplan, & Detlef Schoder

The Role of Cognition and Affect in the Formation of Customer Satisfaction: A Dynamic Perspective 21 Christian Homburg, Nicole Koschate, & Wayne D. Hoyer

Scents and Sensibility: When Do (In)Congruent Ambient Scents Influence Product Evaluations? 32 Anick Bosmans

From Fear to Loathing? How Emotion Influences the Evaluation and Early Use of Innovations 44 Stacy L. Wood & C. Page Moreau

Are All Smiles Created Equal? How Emotional Contagion and Emotional Labor Affect Service Relationships 58 Thorsten Hennig-Thurau, Markus Groth, Michael Paul, & Dwayne D. Gremler

Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue 74 Yong Liu

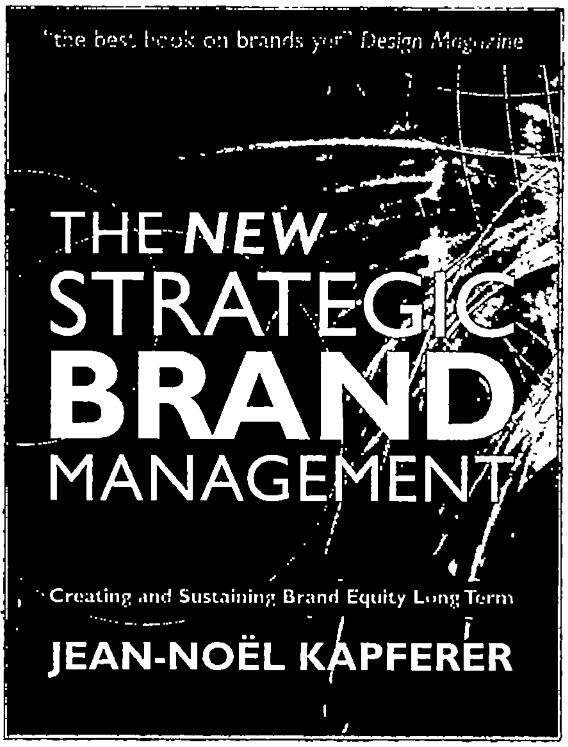
Friends, Businesspeople, and Relationship Roles: A Conceptual Framework and a Research Agenda 90 Jan B. Heide & Kenneth H. Wathne

Human Brands: Investigating Antecedents to Consumers' Strong Attachments to Celebrities 104 Matthew Thomson

Dual Distribution and Intangible Firm Value: Franchising in Restaurant Chains 120 Raji Srinivasan

How Country Characteristics Affect the Perceived Value of Web Sites 136 Jan-Benedict E.M. Steenkamp & Inge Geyskens

One of the cornerstones of brand management in MBA programs today


"the best book on brands yet" Design Magazine
THE NEW
STRATEGIC
BRAND
MANAGEMENT
Creating and Sustaining Brand Equity Long Term
JEAN-NOËL KAPFERER

PAPER, 520 PP; 0 7494 4283 2, $59.95

☐ New models and benchmarks for building brands and businesses
☐ New approach to brand identity and positioning, brand architecture and multi-brand portfolios
☐ Addresses key issues, including:
  – *Building brands in saturated markets*
  – *Managing retail brands*
  – *Energizing brand value chains*
  – *Brand extensions*
  – *Brand innovation*
  – *Launching brands*
  – *Globalizing existing brands*
  – *Revitalizing aging brands*

JN Kapferer (HEC, Ph.D Northwestern University) is considered one of world's top branding experts. He has been a consultant for European, Asian and American companies and teaches advanced executive development seminars worldwide. In 2004, he received the AMA's S. Taumer Cavusgil Award.

KOGAN PAGE

For more details, visit **www.koganpage.com** *or contact us for an inspection copy:*

525 South 4th Street, Suite 241 • Philadelphia, PA 19147
Telephone: 215-928-9112 or 215-525-3251
Fax 215-928-9113 • E-mail: info@koganpage.com

Journal of
Marketing

Vol. 70, No. 3

AMERICAN MARKETING ASSOCIATION

Calcutta University Library
18·8·06
ALIPORE CAMPUS

EDITOR
Roland T. Rust
Robert H. Smith School of Business
3451 Van Munching Hall
University of Maryland
College Park, MD 20742-1815

ASSISTANT TO THE EDITOR
Marti Tiedeman
E-mail: jom@ama.org
Tel: (312) 542-9024

Millicent Locke
University of Maryland

WEB EDITOR
Siva K. Balasubramanian
Southern Illinois University

MARKETING LITERATURE REVIEW
Myron Leonard
Western Carolina University

BOOK REVIEWS
Ramarao Desiraju
University of Central Florida

VICE PRESIDENT—PUBLICATIONS
Russell S. Winer
New York University

CHIEF EXECUTIVE OFFICER
Dennis Dunlap

GROUP PUBLISHER
Jack Hollfelder

EDITORIAL STAFF
Phone: (312) 542-9000
Fax: (312) 542-9001

MANAGING EDITOR
Francesca Van Gorp Cooley

TECHNICAL EDITOR
Andy Seagram

GRAPHICS EDITOR
Mary Loye

Advertising Staff
Fax: (312) 922-3763

Richard Ballschmiede, Advertising
(312) 542-9076

Sally Schmitz, Production
(312) 542-9038

JM Web site:
Submissions: www.marketingjournals.org/JM
Main page: www.marketingpower.com/JM

The *Journal of Marketing* (ISSN 0022-2429) is published quarterly (January, April, July, and October). *Annual Subscription Rates*: Subscriptions are $95 for individuals in the United States and Canada and $120 in other countries. Library, school, and business subscriptions are $235 in the United States and Canada and $265 in other countries. Subscriptions for AMA members are $55. Prices are for one year. Single copy rate: $50 individual, $60 corporate in the United States and Canada, add $5 in other countries. Prepayment is required in U.S. dollars or the equivalent.

Published by the Publications Group of the American Marketing Association, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266. Periodicals postage paid at Chicago, IL and additional mailing offices. POSTMASTER: Send address changes to the *Journal of Marketing*, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266, USA.

•**MANUSCRIPTS**: Electronic submissions must be made via the *JM* Web site at http://www.marketingjournals.org. Paper submissions are not accepted. Editorial policies, review board members, Tables of Contents for *JM*, and so on, are available on the *JM* Web site at http://www.marketingpower.com/jm.

•**ADVERTISERS AND ADVERTISING AGENCIES** assume liability for all content (including text, representations, and illustrations) of advertisements published, and also assume responsibility for any claims arising therefrom made against the publisher. The right is reserved to reject any advertisement.

•**SUBSCRIPTIONS**: Members: Send to American Marketing Association, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266 (1-800-AMA-1150 or 312/542-9000), or e-mail to info@ama.org. Nonmembers: Send to American Marketing Association, P.O. Box 11806, Birmingham, AL 35202, 1-800-633-4931, 205-995-1567, Fax 205-995-1588, e-mail: amasubs@ebsco.com. Canadian residents add 7% GST (GST registration # 12747852).

•**PERMISSIONS AND REPRINT INFORMATION:** Without written permission from the AMA, any copying or reprinting (except by authors reprinting their own work) is prohibited. Requests for permission to reprint—such as copying for general distribution, advertising, or promotional purpose; for creating new collective works; or for resale—should be submitted in writing to permissions@ama.org. Bulk reprints are available by contacting Tamara Groft, Sheridan Reprints, (800) 352-2210. Permission is granted by AMA for those registered with the Copyright Clearance Center (CCC), 222 Rosewood Dr., Danvers, MA 01923, to photocopy any article herein for a flat fee. Contact CCC with ISSN 0022-2429, volume, and first and last page numbers of each article copied.

The views expressed in *JM* are not necessarily those of the editorial staff or the publisher. **All articles in this journal, unless otherwise noted, have undergone a blind review process.**

Copyright © 2006, American Marketing Association. All rights reserved

Editorial Review Board

ROHINI AHLUWALIA
University of Minnesota

KUSUM L AILAWADI
Dartmouth College

TIM AMBLER
London Business School

EUGENE W ANDERSON
University of Michigan

TOR WALLIN ANDREASSEN
Norwegian School of Management

ASIM ANSARI
Columbia University

ERIC J ARNOULD
University of Arizona

BARRY L BAYUS
University of North Carolina

PAUL D. BERGER
Boston University

PAUL N BLOOM
University of North Carolina

LISA E. BOLTON
University of Pennsylvania

RUTH N BOLTON
Arizona State University

ROBERT BORDLEY
General Motors Corporation

DOUGLAS BOWMAN
Emory University

RODERICK J BRODIE
University of Auckland

DIPANKAR CHAKRAVARTI
University of Colorado

PIERRE CHANDON
INSEAD

PRADEEP K. CHINTAGUNTA
University of Chicago

KEVIN CLANCY
Copernicus Marketing Consulting

PETER J DANAHER
University of Auckland

JOHN A DEIGHTON
Harvard University

MARNIK G DEKIMPE
Catholic University Leuven

WAYNE S DESARBO
Pennsylvania State University

RAVI DHAR
Yale University

NAVEEN DONTHU
Georgia State University

TULIN ERDEM
University of California, Berkeley

PETER S FADER
University of Pennsylvania

GARY L FRAZIER
University of Southern California

PETER N GOLDER
New York University

RAJDEEP GREWAL
Pennsylvania State University

ABBIE GRIFFIN
University of Illinois, Urbana-Champaign

SUNIL GUPTA
Columbia University

JOHN R HAUSER
Massachusetts Institute of Technology

JAN B. HEIDE
University of Wisconsin, Madison

CHRISTIAN HOMBURG
University of Mannheim

WAYNE D. HOYER
University of Texas at Austin

G TOMAS M HULT
Michigan State University

J. JEFFREY INMAN
University of Pittsburgh

GANESH IYER
University of California, Berkeley

CHRIS JANISZEWSKI
University of Florida

SANDY D JAP
Emory University

BERNARD J JAWORSKI
Monitor Group

MICHAEL D JOHNSON
University of Michigan

PK KANNAN
University of Maryland

TIMOTHY L. KEININGHAM
Ipsos Loyalty

AMNA KIRMANI
University of Maryland

AJAY K. KOHLI
Emory University

ROBERT V. KOZINETS
York University

V. KUMAR
University of Connecticut

DONALD R. LEHMANN
Columbia University

KATHERINE N. LEMON
Boston College

MARY FRANCES LUCE
Duke University

ROBERT F. LUSCH
University of Arizona

JOHN G. LYNCH JR.
Duke University

DEBORAH J. MACINNIS
University of Southern California

VIJAY MAHAJAN
University of Texas at Austin

CARL F. MELA
Duke University

VIKAS MITTAL
University of Pittsburgh

CHRISTINE MOORMAN
Duke University

PAGE MOREAU
University of Colorado

SCOTT A. NESLIN
Dartmouth College

STEPHEN M. NOWLIS
Arizona State University

JOSEPH C. NUNES
University of Southern California

A. PARASURAMAN
University of Miami

WILLIAM D. PERREAULT
University of North Carolina

F.G M PIETERS
Tilburg University

JAIDEEP PRABHU
Imperial College

ARVIND RANGASWAMY
Pennsylvania State University

BRIAN T. RATCHFORD
University of Texas at Dallas

AMERICUS REED
University of Pennsylvania

WERNER REINARTZ
INSEAD

STEVEN M. SHUGAN
University of Florida

ITAMAR SIMONSON
Stanford University

DILIP SOMAN
University of Toronto

JOYDEEP SRIVASTAVA
University of Maryland

JAN-BENEDICT E M STEENKAMP
Tilburg University

DAVID W. STEWART
University of Southern California

STEFAN STREMERSCH
Erasmus University Rotterdam

BAOHUNG SUN
Carnegie Mellon University

GERARD J. TELLIS
University of Southern California

STIJN VAN OSSELAER
Erasmus University Rotterdam

P. RAJAN VARADARAJAN
Texas A&M University

STEPHEN L VARGO
University of Hawaii, Manoa

PETER VERHOEF
University of Groningen

MICHEL WEDEL
University of Maryland

BARTON A WEITZ
University of Florida

RUSSELL S. WINER
New York University

ANTHONY ZAHORIK
Burke Institute

VALARIE A. ZEITHAML
University of North Carolina, Chapel Hill

Z. JOHN ZHANG
University of Pennsylvania

From the Editor: The Maturation of Marketing as an Academic Discipline

Journal of Marketing (*JM*) is now in the middle of its 70th volume, and the changes in the journal reflect the changes in the marketing field. In short, this is a transitional period for the marketing discipline, one that reflects the maturation of marketing thought and a growing self-confidence in the discipline's ability to create important new knowledge on its own.

Marketing as an Applied Discipline

Unlike many of the core disciplines from which marketing borrows (e.g., economics, psychology, anthropology), marketing arose as a discipline taught in business schools, primarily for the purpose of helping marketing managers do their jobs better. This applied orientation for the field ties academic thought in marketing to the real world, either directly or indirectly. The business world benefits tremendously from academic publications in journals such as *JM*, as a visit to any sophisticated practitioner conference, such as the American Marketing Association (AMA) Advanced Research Techniques Forum, the AMA Frontiers in Service Conference, or any Marketing Science Institute conference, will quickly confirm. This applied orientation also has its advantages for marketing academics, in that the perceived usefulness of their work has led to higher salaries compared with professors in the humanities or the pure sciences.

Science Versus Technology

The applied nature of marketing has created some internal conflicts in the academic field because much of the academic culture of marketing has been borrowed from core disciplines in the social sciences, notably, from psychology and economics. The pure science culture tends to treat knowledge as inherently good and to disregard (or even disdain) application. This culture pervades the social sciences, in which application of scientific knowledge is largely considered outside the academic domain. Such a dichotomy is also found in the physical sciences, in which fundamental scientific discovery is in conflict with, but is also symbiotic with, technological application. For example, applications in solid state physics in a physics department may make technological advances possible in an engineering department. Similarly, mathematical advances in a mathematics department may create the fundamental basis for new computational technologies in a computer science department. Viewed in these terms, marketing departments should perhaps be most similar to departments that create more applied technologies, such as engineering or computer science.

Nevertheless, the actual evolution of marketing departments has resulted in a science–technology dichotomy, often existing within the same department, and certainly both points of view coexist within the field of marketing itself. This has led to certain misunderstandings, with science-oriented researchers looking down on technologically oriented researchers, and vice versa.

Legitimate Purposes for Research in Marketing

The science–technology dichotomy implies two main legitimate purposes for research in marketing. The first is research that advances fundamental knowledge in marketing and furthers scientific thought. Winners of *JM*'s Harold H. Maynard Award are sometimes articles of this type (e.g., MacInnis and De Mello 2005). Most of the consumer behavior research that is published in the field of marketing is of this nature. In some journals (e.g., *Journal of Consumer Research*), concern for the practical application of scientific knowledge has traditionally been deemphasized.

The second legitimate purpose for research in marketing is research that uses existing fundamental scientific knowledge to advance practical interests (usually managerial interests but also consumer interests or public policy interests). Many of the quantitative decision models in marketing are of this type. For this type of research, there is (and should be) less emphasis on advancing scientific thought and more interest in inventing something new that is useful. Some examples of this approach are recent winners of *JM*'s Marketing Science Institute/H. Paul Root Award (Reinartz, Thomas, and Kumar 2005; Rust, Lemon, and Zeithaml 2004). Such articles often appear in the more quantitatively oriented journals, such as *Journal of Marketing Research* and *Marketing Science.*

To some degree, *JM* is a bridge between these two purposes. Many of its articles advance scientific thought in marketing, but there is always attention to how the work is meaningful to stakeholders in marketing.

Homegrown Theory

How to deal with theory is one of the major issues that the marketing field must address as it matures. In the scientific paradigm, theory is essential to provide generalizability and efficient ways of understanding. Too often, this has been understood as requiring "basic discipline" theory, meaning that a theory must be fundamental to psychology, economics, or some other basic discipline. Often, this has been interpreted to mean that research in marketing that does not borrow fundamental theory from one of the basic disciplines is "atheoretical." Marketing will only become mature as a field when it is realized that it is also possible to have theory that is indigenous to marketing—that is, theory that

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

does not necessarily result in a fundamental advance in one of the basic disciplines. There are already some good examples of this in the literature (e.g., Shugan and Xie 2000; Vargo and Lusch 2004).

Reviewers and editors must learn to be accepting of such theory. Too often, I have observed reviewers react to new, marketing-related theory by saying that it is "ad hoc." By ad hoc, they really mean that it has been created by the author rather than borrowed from a core discipline. In marketing, there must be more acceptance of homegrown theory that is developed on the basis of problems and issues that are not only specific to marketing but also generalizable within marketing. For example, it should be acceptable to develop behavioral theory that is relevant to marketing but not particularly fundamental to psychology. Likewise, it should be acceptable to develop theory in the form of mathematical models that are useful to marketing but not necessarily of interest to economics.

The Changing Environment

The maturation of the marketing discipline also means keeping up with the technological and economic environment. As technology advances, the economy changes in predictable ways that affect how marketing should be practiced and studied. Technology—particularly, information technology—is making it easier to communicate with customers, store information about customers, and process this information to provide better service to customers. The result is that forming and maintaining relationships with customers on a large-scale basis is easier than it has ever been. This results in an expansion of the importance of service in the economy. Because technology development is a one-way street—it is always moving forward—this expansion of the importance of service is certain to continue, and the concepts of service marketing and relationship marketing will become increasingly central to the field.

Technology has also radically changed the way research is conducted, and it will continue to do so. It is clear, though perhaps easy to forget, how much more computational research in marketing became in the 1960s and 1970s when computers first became widely usable for research. Today, almost every doctoral student takes for granted the ability to use high-speed computing capability for data analysis. This trend will continue, and it is safe to predict that future research in marketing will become even more computationally focused, with likely advances coming from approaches that are highly computationally intensive. This has already been observed in the past ten years or so as Bayesian methods have become mainstream, and in the next ten years, large-scale simulations will likely become widely used as a research approach.

The effect of technology on communication technologies also results in globalization of the economy. Again, this is a one-way street, and the economy will become increasingly global over time. Marketing practice is increasingly global, and the same trend is reflected in marketing academics. To mature fully as a discipline, marketing must become fully global, and the related journals, conferences, and research papers must reflect this. As such, *JM* has tried to lead in this effort by expanding the international diversity of the editorial board. Another model of internationalization is the 2006 AMA Frontiers in Service Conference, which will be held in Brisbane, Australia, and will feature presenters from a record 29 countries. To further *JM*'s international agenda, as editor, I have made it a point to travel internationally whenever possible and to encourage international submissions.

Summary

After 70-plus years as an academic discipline, the marketing field has come a long way, but it needs to make further progress to be truly mature. We need to recognize our purpose as an applied discipline and respect and tolerate the conflicting value systems inherent in the science and technology paradigms. We need to recognize that advancing science is not the only legitimate purpose for an academic article in marketing, nor is advancing the technology of marketing by inventing practical new models or devising new procedures. Both kinds of advances should be encouraged. Crucially, we need to be confident enough as a discipline to create and accept indigenous theory that is based on marketing issues rather than that of a core discipline. We need to modernize and update our field by reflecting the changes created by technological advances, which means an increasing focus on service and relationships, greater use of computational methods, and increasing globalization. The mature marketing discipline will be tolerant, confident, global, technologically savvy, and unashamed to be primarily about marketing.

ROLAND T. RUST
Editor

REFERENCES

MacInnis, Deborah J. and Gustavo E de Mello (2005), "The Concept of Hope and Its Relevance to Product Evaluation and Choice," *Journal of Marketing*, 69 (January), 1–14.

Reinartz, Werner, Jacquelyn S. Thomas, and V. Kumar (2005), "Balancing Acquisition and Retention Resources to Maximize Customer Profitability," *Journal of Marketing*, 69 (January), 63–79.

Rust, Roland T., Katherine N. Lemon, and Valarie A. Zeithaml (2004), "Return on Marketing: Using Customer Equity to Focus Marketing Strategy," *Journal of Marketing*, 68 (January), 109–127.

Shugan, Steven M. and Jinhong Xie (2000), "Advance Pricing of Services and Other Implications of Separating Purchase and Consumption," *Journal of Service Research*, 2 (3), 227–39.

Vargo, Stephen L. and Robert F. Lusch (2004), "Evolving to a New Dominant Logic for Marketing," *Journal of Marketing*, 68 (January), 1–17.

From the Web Editor: Improvements to the *Journal of Marketing* Web Site

I am pleased to highlight several features that have been recently added to the *Journal of Marketing* (*JM*) Web site (http://www.marketingpower.com/jm). These improvements aim to increase the salience of *JM* in readers' professional careers, encourage greater interaction between *JM* readers and authors, and share new information about the journal. I take this opportunity to thank American Marketing Association (AMA) colleagues Chris Leporini, Francesca Van Gorp Cooley, and James Heckman for their help in facilitating these improvements and Charles Hofacker for promoting the improvements in the Academic Resource Center section of the AMA Web site. Please try the following features that are currently available at the *JM* Web site.

Current Features

Make JM *Your Default Home Page*

Readers are only a click away from making *JM* a prominent part of their professional life. A "Make *JM* Your Home Page" button on the main Web site (and linked pages) helps set *JM* as the default home page for Internet Explorer users. We plan to extend this feature to other Internet browsers.

Visit the "Book Review" Section

This section has transitioned to the Web site under the leadership of the *JM* Book Review Editor Ramarao Desiraju. This process has opened opportunities for new features that Ram has planned, such as cross-linking book reviews to relevant content at other AMA-related sites and incorporating readers' inputs about the content of book reviews. Overall, this section appears exciting.

Blog on Specific JM *Articles*

Blogging combines the immediacy of e-mail with the visibility and permanence of a public, informed, and ongoing online dialogue that is vital for community-building efforts. To encourage blogging by *JM* authors and readers, a URL tag line will accompany each article beginning with *JM*'s July 2006 issue. The broad goal is to stimulate a continuous public discourse on each *JM* publication. In particular, blog-stimulated interactions between marketing practitioners and academics are likely to enhance the relevance of published research; they should also contribute new and interesting insights that advance the discipline. Authors of all *JM* articles now have the opportunity to prepare an executive summary along the following lines: a list of short questions addressed in that article, a few sentences that answer these questions by showcasing key findings, and, finally, a list of short issues that the article offers for further research. The executive summary is posted in an article-specific URL and serves as a starting point for open-ended discussion about that article through blogs that participants contribute. The format of the executive summary intentionally differs from the standard abstract of *JM* articles to encourage participation from and interactions with marketing managers. I urge all authors and readers to participate actively in this blogging forum. Related future plans include preparing frequently asked questions and search/retrieval features that support blogging.

Check Out the List of "Most-Cited JM *Articles"*

The *JM* Web site showcases a list of its top ten most-cited articles; this list is based on information supplied by Google Scholar. These rankings are currently anchored to the 2000–2005 period and represent a snapshot as of the date indicated. The rankings are updated each quarter, and Web site visitors can access ranking information from previous quarters. Additional plans for future improvements include the identification of "classic" *JM* articles from each decade (from the 1960s through the 1990s) and new metrics to assess the popularity of *JM* articles (e.g., the number of downloads of specific *JM* articles through the "pay-per-view" option that AMA provides).

Get Acquainted with JM *Statistics*

The Web site includes pie charts that depict various subscriber types for both members and nonmembers. In the future, we plan to include graphical information on citation and impact statistics for the journal.

Future Features

I plan to add a data/code repository at the *JM* Web site, a feature that supports both replication research and graduate student learning. For example, researchers interested in replicating the work in a *JM* article may benefit from the ability to download the data used therein. Similarly, marketing academics who teach doctoral seminars (or rigorous MBA classes) may develop empirically focused student assignments that are structured around specific journal articles. Authors of published *JM* articles can help by contributing a copy of their data sets/program codes to a repository housed within the *JM* Web site. This approach requires sensitivity to protecting the right of participating *JM* authors to use their data in their own future publications. Thus, *JM* plans to restrict downloads of repository content to researchers who submit to a rigorous registration/authentication process and who commit to using the extracted data/code only for research replications (not research extensions).

I hope you find the current (and planned) features at the *JM* Web site useful. Ultimately, the *JM* Web site is a

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

resource for the entire *JM* community, so meaningful improvements to the Web site critically hinge on readers' suggestions and inputs. Please comment on the aspects you like (and those you do not like), and take a few moments to propose new features that will help you. I especially welcome your thoughts on how the journal's Web site can better serve you (siva.jmwebsite@gmail.com).

SIVA K. BALASUBRAMANIAN
Web Editor

Michael Haenlein, Andreas M. Kaplan, & Detlef Schoder

Valuing the Real Option of Abandoning Unprofitable Customers When Calculating Customer Lifetime Value

In recent years, several authors have developed models that focus on the allocation of scarce marketing resources based on customer lifetime value (CLV). These approaches use CLV to develop a rank order of customers and recommend devoting more resources to customers with higher ranks. However, it has been discussed in the literature that a simple net present value analysis may not reflect the value of the flexibility to make such decisions. Therefore, some authors recommend the use of a real-options analysis in certain situations. Building on this stream of research and using the case of the option to abandon unprofitable customers, this article proposes an approach that combines real-options analysis and CLV; this approach explicitly values the seller's flexibility to abandon unprofitable customers. Using a combination of examples, empirical analysis, and Monte Carlo simulations, the authors provide evidence that the divergence between CLV that includes and CLV that excludes option value can be substantial and may not be the same for all customers. Therefore, the authors conclude that a distribution of customers based on CLV can change when option value is included. Thus, using CLV as a basis for marketing decisions but not including the value of the option to make such decisions a priori when calculating CLV can lead to an overall biased result.

Approximately 30 years ago, Kotler (1974, p. 24) defined long-term customer profitability as "the present value of the future profit stream expected over a given time horizon of transacting with the customer." Since then, customer lifetime value (CLV) has contributed to solving various problems, from mailing decisions in the catalog sales industry (Bitran and Mondschein 1996) to decisions related to mergers and acquisitions (Selden and Colvin 2003) to questions about company valuation (Gupta, Lehmann, and Stuart 2004). In recent years, several authors have developed models that focus on the allocation of scarce marketing resources based on CLV (e.g., Reinartz, Thomas, and Kumar 2005; Rust, Lemon, and Zeithaml 2004; Venkatesan and Kumar 2004). These approaches use CLV to develop a rank order of customers and recommend devoting more resources to customers with higher ranks. Some authors go so far as to state that customers with low ranks (especially when these show negative CLV) should be abandoned completely to increase overall profitability of the customer base (Zeithaml, Rust, and Lemon 2001).

However, it has been discussed in the literature that the simple net present value (NPV) analysis may not reflect the value of the flexibility to make such decisions. Therefore, some authors recommend the use of a real-options analysis in certain situations (e.g., Amram and Kulatilaka 1999; Dixit and Pindyck 1995; Pindyck 1991). To the best of our knowledge, real-options analysis has been investigated only twice in the context of CLV, namely, in a publication by Levett and colleagues (1999) and a working paper by Hogan and Hibbard (2002). Levett and colleagues propose that the relationship between a buyer and a seller can be expressed as a set of options. Because the customer has the option to choose between buying and not buying at any point in time, the authors determine CLV using a binomial option valuation model. Hogan and Hibbard develop a real-options-based framework for assessing the value of business relationships. They state that the value of any relationship is influenced by the core benefit value and the value of two options: a switching option (i.e., the flexibility to switch to another seller or buyer) and a growth option (i.e., the freedom to expand the relationship beyond its current scope). Thus, relationships can be categorized on the basis of a three-dimensional typology, depending on the magnitude of each of these value components.

In this article, we propose an alternative approach to combining real-options analysis and CLV that explicitly values the seller's flexibility to abandon unprofitable customers. Traditional CLV calculations assume that a company will continue to deal with unprofitable customers, and indeed, a company may do so, in which case regular CLV calculations are appropriate. However, when a firm discon-

---

Michael Haenlein is Assistant Professor of Marketing, Department of Marketing, ESCP-EAP European School of Management (e-mail: haenlein@escp-eap.net). Andreas M. Kaplan is a doctoral candidate (e-mail: andreas.kaplan@uni-koeln.de), and Detlef Schoder is a full professor (e-mail: schoder@wim.uni-koeln.de), Department of Information Systems and Information Management, University of Cologne.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 70 (July 2006), 5–20

tinues unprofitable customers because it believes that the customer relationship cannot become profitable in future periods, we propose a modification of standard CLV calculations. In this sense, our work differs from that of Levett and colleagues (1999) because we focus on modeling the seller's options, whereas they concentrate on the buyer's options and build on the assumption that the seller will always sell the product when the buyer is willing to purchase it. Our article is also different from Hogan and Hibbard's (2002) work because we develop a specific metric (i.e., CLV) to quantify the value of a specific type of market-based asset (i.e., customer relationships), whereas their approach is conceptual and broader in scope. Because our results provide evidence that the divergence between CLV that includes and CLV that excludes option value can be substantial and may not be the same for all customers, we conclude that a distribution of customers based on CLV can change when option value is included. Therefore, using CLV as a basis for marketing decisions but not including the value of the option to make such decisions a priori when calculating CLV can lead to an overall biased result.

Take the mobile phone industry as an example. Because of the absence of indicators of customer quality at the time of acquisition, mobile phone operators usually attribute the same customer acquisition cost (mainly consisting of handset subsidies) to all customers acquired through the same acquisition channel. To prioritize acquisition channels in the allocation of a limited total budget, the CLV of an "average" customer acquired through any channel is calculated. Because customers expect to replace their handset at regular time intervals, this CLV includes the expected cost of future upgrades. In reality, however, companies upgrade customers only at the end of their acquisition contract if the revenues generated until that point indicate sufficiently good quality to justify a reinvestment in the relationship. At this point, revenue that the customer generates is a much better predictor of future customer quality than simply the information about the acquisition channel. Therefore, mobile phone operators keep the freedom to exercise their option not to upgrade when prior revenues are unfavorable. This usually equates to an abandonment decision because a customer who does not receive a desired upgrade is likely to leave. However, CLV does not account for this flexibility, because it assumes a fixed expected cost of future upgrades for all customers acquired through the same channel. Thus, the value of the flexibility not to upgrade low-value customers is not included in the calculation of CLV. This makes the traditional CLV measure inconsistent with actual behavior. Therefore, using CLV determined this way when allocating a limited total customer acquisition budget across acquisition channels is likely to lead to a bias.

The remainder of this article is structured as follows: In the next section, we introduce real-options theory by showing how this valuation technique can overcome some limitations of simple NPV analysis and how it can be applied in a marketing context. We then use two examples to illustrate how the value of the option to abandon unprofitable customers can be calculated. We go on to validate our approach empirically using data provided by the Direct Marketing Educational Foundation (DMEF) to estimate the divergence between CLV that includes and CLV that excludes option value. Subsequently, we conduct a Monte Carlo simulation to identify factors that influence this divergence. The article concludes with a discussion of the theoretical and managerial implications of our findings, as well as the limitations of our approach and areas of further research.

Real-Options Theory

Authors in the area of strategic investment decision making have noted that simple NPV analysis is associated with several implementation and conceptual problems. Although some of these problems can be cured by modifications of this approach that have been discussed in the literature (for a detailed list and explanation, see Adler 2000), a shortcoming remains (e.g., Amram and Kulatilaka 1999; Dixit and Pindyck 1995): NPV analysis assumes that *one* specific value for some kind of metric (e.g., cash flow) can be defined for each period within the time horizon of the analysis. Although it is possible to account for future uncertainty by specifying more than one state of nature per period, different values of this metric must be combined into one expected value before NPV can be calculated. Therefore, simple NPV analysis builds on the assumption that a project can be completely planned beforehand and ignores the project's potential to create future options that are not foreseeable based on the current stock of knowledge. Such options could include the freedom to abandon the project in case of unfavorable conditions or to extend its scope in case of favorable ones. However, it is a characteristic of options to show an asymmetric behavior in net payoffs because they always represent the right but not the obligation to carry out a future action. Thus, because NPV analysis ignores these options, it often underestimates the true value of an investment (e.g., Benninga and Tolkowsky 2002). The literature on the strategic management–finance interface that has emerged around the valuation of such (real) options is extensive, whereas to our knowledge, there are only four studies that discuss this approach in marketing (Finch, Becherer, and Casavant 1998; Hogan and Hibbard 2002; Levett et al. 1999; Slater, Reddy, and Zwirlein 1998). Nevertheless, the application of a real-options approach to customer acquisition/retention analysis using CLV seems to be particularly appropriate. These activities usually create future options because the decision to acquire or reinvest in a customer always creates the choice of further investment in the relationship in subsequent periods.

An essential step after identifying the existence of a real option is its valuation. For this task, there are three different approaches that can be used: Black and Scholes's (1973; hereinafter, the Black–Scholes model) model, contingent claims valuation (Cox, Ross, and Rubinstein 1979), and dynamic programming (for an application to real-options valuation, see Copeland and Tufano 2004). The Black–Scholes model has the advantage of providing a closed-form solution for option value. It has previously been applied in the context of real-options analysis (e.g., Luehrman 1998), but its drawback is that it is computationally demanding. Contingent claims valuation is much easier to apply than the Black–Scholes model and uses less stringent assumptions, though it includes the Black–Scholes formula as a limiting case. It has also been used to value real options (e.g., Sachdeva and Vandenberg 1993), but it builds on the assumption of arbitrage-free markets and thus assumes that the real option is tradable. This assumption may be questionable in the context of CLV calculations. In addition, as Devinney and Stewart (1988) note, models that have their roots in financial markets cannot always be applied to product markets, because this may violate key assumptions underlying them. Therefore, we do not recommend either of these two approaches for valuing real options in a marketing context; rather, we suggest the use of a third method, the dynamic programming approach.

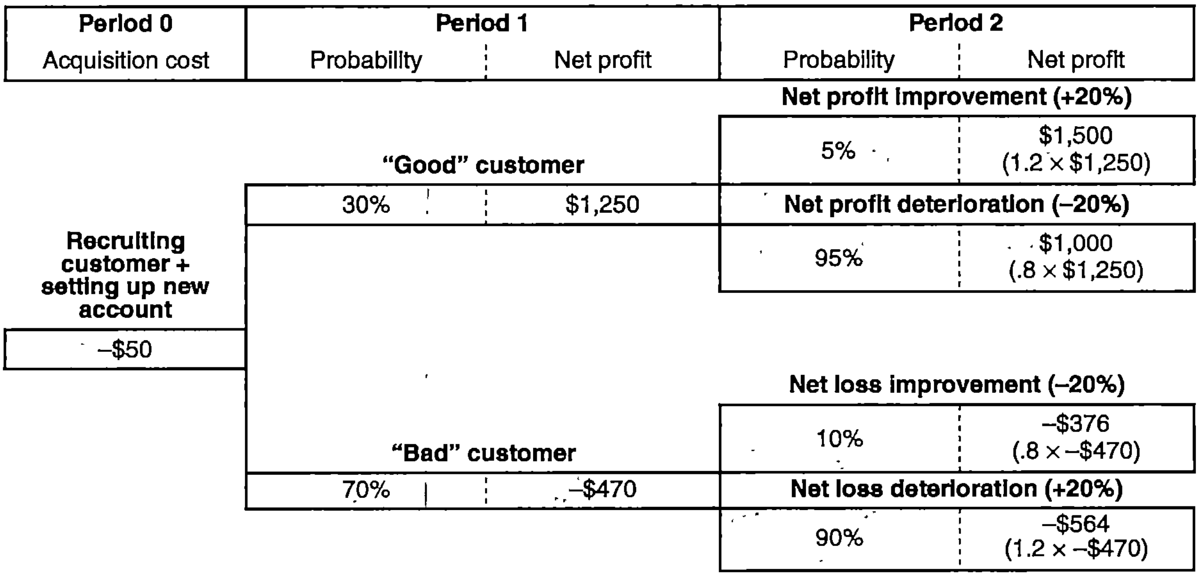
Dynamic programming is a class of solution methods for solving sequential decision problems. It builds on the assumption that "the optimum way of getting from a particular state to an end point is not a function of the path taken to that state" (Elton and Gruber 1971, p. 480). This is often also referred to as Bellman's principle of optimality. Thus, the optimal decision at any time t can be identified by maximizing the sum of the paybacks from following a particular policy at time t plus the expected paybacks from following an optimal policy from time t + 1 onward. In the context of real-options analysis, dynamic programming helps examine options from more of a decision theory point of view. It is not based on the same assumptions underlying financial option valuation. Nevertheless, despite this fundamental difference, the contingent claim value of an investment can also be identified by the dynamic programming approach. However, this is true only after parameter calibration (Knudsen, Meister, and Zervos 1999) because the dynamic programming method assumes an exogenously determined discount rate, whereas contingent claims valuation determines a risk-adjusted discount rate endogenously.

In the following section, we use two examples to show how dynamic programming can be applied to valuing a seller's abandonment option to terminate business with a customer who has proved to be unprofitable. We propose a real-options-based approach to CLV, which explicitly values the option to abandon unprofitable customers.

Examples

Imagine a marketing manager who must choose between implementing one of two alternative advertising campaigns (A and B). With Campaign A, the company can attract "good" customers with a probability of 30% and "bad" ones with a probability of 70%. A good client will generate a net profit of $1,250 in Period 1, whereas a bad one is assumed to result in a net loss of –$470 in the same period. For both types of clients, this net contribution can either improve or deteriorate by 20% in Period 2, whereas the likelihood of an improvement will be 5% for a good and 10% for a bad customer. In this context, the term improvement is viewed from a company perspective and represents either a profit increase or a loss decrease (see Figure 1). With Campaign B, only one type of customer, one that generates a constant net profit of $46 in Periods 1 and 2, can be acquired. For both campaigns, the average cost of recruiting the customer and setting up a new account is assumed to be $50, and the discount rate is 15%. Using an average lifetime of two periods, the CLV of a newly acquired customer based on a simple NPV analysis is approximately –$16 for Campaign A and $75 for Campaign B. Because the CLV of Campaign B is greater than that of Campaign A and greater than the acquisition cost, the decision would be in favor of Campaign B.[1]

FIGURE 1
Expected Net Contribution of Customers Acquired Through Campaign A (Example 1)



---

[1] Although the example is an artificial one, the key numbers can be considered realistic for the special case of the retail banking industry. The net contribution of a good and bad customer in Campaign A is determined on the basis of an expected revenue of $1,600 and $230 and the expected cost of $350 and $700, respec-

However, as we noted previously, NPV analysis does not account for the flexibility of the marketing manager after the end of Period 1. At that point, the marketing manager knows (e.g., on the basis of experience from comparable campaigns already carried out by the company) that a customer who proved to be bad in Period 1 will generate an expected loss in Period 2. Therefore, the marketing manager has the option to behave asymmetrically in Period 2; that is, in the case of a good customer, the client relationship will continue, and in the case of a bad one, it will be abandoned. This asymmetrical behavior protects the manager from the downside risk to realize the loss of a bad customer in Period 2. Simple NPV analysis does not account for this option because only one value for the expected net contribution of a customer in Period 2 needs to be determined. We assume that the marketing manager decides to terminate the client relationship with a bad customer at the end of Period 1 and that this abandonment does not cause additional cost. In this case, the lifetime value of a bad customer increases from –$821 to –$409, which increases the expected CLV from –$16 to $273. On the basis of this increase of $289, which can be interpreted as the value of the abandonment option, Campaign A now seems far more attractive than Campaign B.

The previous example builds on a simple situation and attempts to explain the basic idea behind our approach in more detail. Now, we apply our approach to a more complex and, thus, more realistic situation.

Imagine that a marketing manager is faced with the decision to acquire a new customer, Mrs. Smith. Again, we assume that this acquisition costs $50. In every period after acquisition, Mrs. Smith can purchase a high-margin product, a low-margin product, or no product at all. We assume that a high-margin product generates a revenue of $200 and that a low-margin product generates a revenue of $100. In every period, the bank invests marketing costs into the relationship with Mrs. Smith. In the first period after acquisition, these costs are assumed to be $100. In every following period, they depend on Mrs. Smith's purchase in the previous period. If she purchases a high-margin product, marketing costs will be $100; if she purchases a low-margin product, marketing costs will drop to $50; and if she makes no purchase, marketing costs will be $25. After every period, the marketing manager has the choice to terminate the relationship with Mrs. Smith. However, we assume that this decision is associated with an abandonment cost of $25. On the basis of previous experience with customers similar to Mrs. Smith in terms of a set of observable characteristics (e.g., acquisition channel, demographics, socioeconomic status), the marketing manager expects a purchasing behavior for the three periods after acquisition (see Figure 2). The discount rate assumed for the analysis is 15%.

Using this information, Mrs. Smith's CLV is $46.51 when option value is excluded. Because this is less than the acquisition cost of $50, the decision would be not to acquire Mrs. Smith. However, this CLV does not account for the option to behave asymmetrically in future periods. After every period, the marketing manager can terminate the relationship with Mrs. Smith depending on the future net contribution streams expected from this relationship. If we apply a dynamic programming approach to this example, there are three points at which abandonment is economically sensible: (1) when Mrs. Smith purchases a high-margin product in the period after acquisition, (2) when she purchases a low-margin product in the period after acquisition and a high-margin product in the next period, and (3) when she makes no purchase in the period after acquisition and purchases a low-margin product in the next period. Including the value of the option to abandon Mrs. Smith at these three points in time leads to a CLV of $52.97, which is greater than the acquisition cost. Thus, when option value is accounted for, the acquisition of Mrs. Smith is a worthwhile strategy.

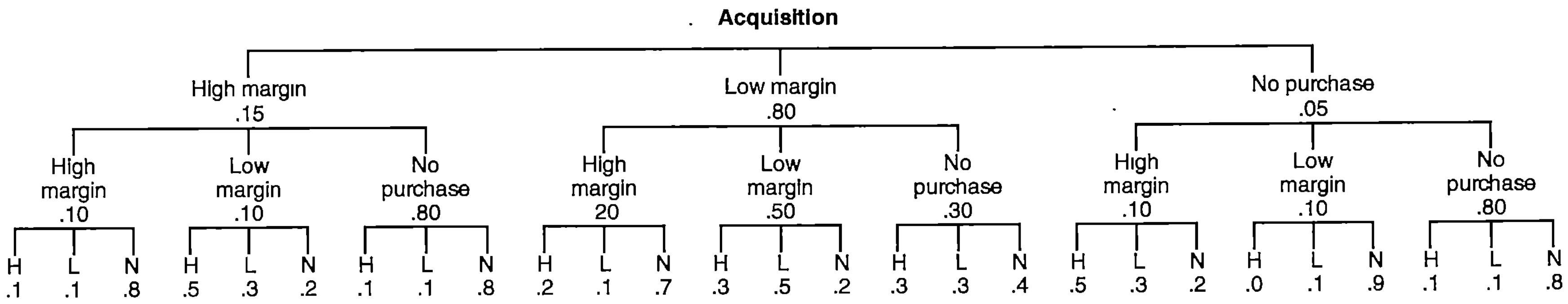
Abandoning Mrs. Smith at these three points makes sense because of a combination of two factors: First, there is a relatively high probability of no purchase in the third period after acquisition, especially when purchases have been carried out in previous periods. This can be explained either by Mrs. Smith's saturation in terms of product consumption or by budget constraints she might be facing. For example, after Mrs. Smith purchases a high-margin product in Periods 1 and 2 after acquisition, the probability that she will not purchase in the following period is 80%. Second, marketing expenditures depend on previous purchasing patterns. This makes a no-purchase scenario highly unattractive when, for example, it is preceded by a high-margin product purchase. There are also other situations in which Mrs. Smith generates a loss; however, in such cases, this loss is less than the abandonment cost of $25, so the abandonment is not a recommendable strategy (given the time frame chosen for analysis). This example shows that in a more complex and realistic setting, the value of the abandonment option requires calculations that can no longer be easily carried out using simple spreadsheet methods. Thus, there is a need to apply more advanced option-pricing approaches.

At this point, it is important to note how real-options analysis differs from two other techniques that have been discussed in the literature: Bayesian updating and elementary decision tree analysis. Bayesian updating is closely linked to real-options analysis in that both techniques deal with the revision of prior beliefs in light of new information. However, there are some specific characteristics of real options, for example, the inherent asymmetry included in their analysis or that the value of a real option increases with both reduced and increased uncertainty. Thus, it is more appropriate to view Bayesian updating as a routine component of real-options analysis than to view both techniques as identical (for a more detailed discussion, see Herath and Park 2001; Miller and Park 2005). With respect

---

tively. This corresponds to the figures given for the most profitable 6% and the least profitable 14% for customers at First Chicago (O'Sullivan 1997, p. 44). The customer acquisition cost of $50 is in line with Reichheld and Sasser's (1990, p. 106) estimations for the credit card industry. We also refer to the work of Berger, Weinberg, and Hanna (2003) for a detailed discussion of how these values can be estimated using data about comparable campaigns already carried out by the company. All CLVs we report in this section exclude acquisition cost.

FIGURE 2
Expected Future Purchasing Probabilities of Mrs. Smith (Example 2)



to the difference between real-options and elementary decision tree analysis, our example could give the impression that real-options analysis is only the folding back of a decision tree. For our example, these two approaches are indeed equal because we voted for the use of dynamic programming as a valuation technique. Nevertheless, this equality is not generalizable to all situations, because other researchers might want to vote for other valuation techniques. For example, the Black–Scholes model, which we apply subsequently in the context of our Monte Carlo simulation, has its roots in the solution of heat-transfer equations in physics and therefore is not related to decision tree analysis.

Empirical Validation

The aforementioned examples, which are built on hypothetical figures, should be viewed only as an illustration of our basic idea. In this section, we focus on empirically validating our approach and on obtaining an estimate of the divergence between CLV that includes and CLV that excludes option value. We use empirical data provided by the DMEF (academic data set three), which consists of 12 years of sanitized purchasing data for more than 100,000 customers of a specialty catalog company. Our empirical validation consists of four steps: First, we estimate the future purchasing behavior (i.e., probability of purchase and dollar amount expected to be spent) for a set of customer segments using the recency, frequency, and monetary value (RFM) approach. Second, we determine costs generated per customer per year. Third, we use these two inputs to estimate the expected future revenue and profit contribution for each customer segment. Fourth, we calculate CLV that includes versus CLV that excludes option value and compare these values on customer base level to obtain an estimate of the divergence between them (for computational details regarding our approach, see the Appendix).

The first step, the estimation of the future purchasing behavior for a set of customer segments, is carried out with the RFM approach. The RFM approach, which Cullinan (1978) first presented and Bauer (1988) subsequently extended, assumes that a customer's future purchase behavior is determined by three purchase history variables: recency, frequency, and monetary value. Recency is the time elapsed since the previous purchase, frequency is the number of purchases made in a specified time period, and monetary value is the dollar sales value of prior purchases (Bauer 1988). Although other, more advanced techniques have been proposed in the literature to model future customer behavior, such as the Pareto/NBD model (Schmittlein, Morrison, and Colombo 1987) and the beta-geometric/NBD framework (Fader, Hardie, and Lee 2005a), we decided to apply the RFM approach in our context. Because the focus of our empirical validation lies in comparing two approaches to determining CLV given an assumed future purchasing behavior, and not in evaluating the predictive accuracy of these methods or in proposing a better method to forecast customer activity, the RFM approach, which is widely known and easy to use, is preferable.

To use the RFM approach, we first split the 12 years of historical purchase data available in our data set into two groups, the first one covering two-thirds (8 years, from t – 12 to t – 5) and the second group covering one-third (4 years, from t – 4 to t – 1). We then determine recency (operationalized as the number of years since the last order) and frequency (measured as the total number of years in which at least one purchase occurred) as of $t_0 = t - 4$. This leads to eight values for recency and frequency. From the total number of 106,284 customers in our data set, 47,617 could be assigned to one of the resulting 36 distinct recency–frequency cells. The remaining 58,667 customers did not make any purchases between t – 5 and t – 12 and were discarded from the analysis.

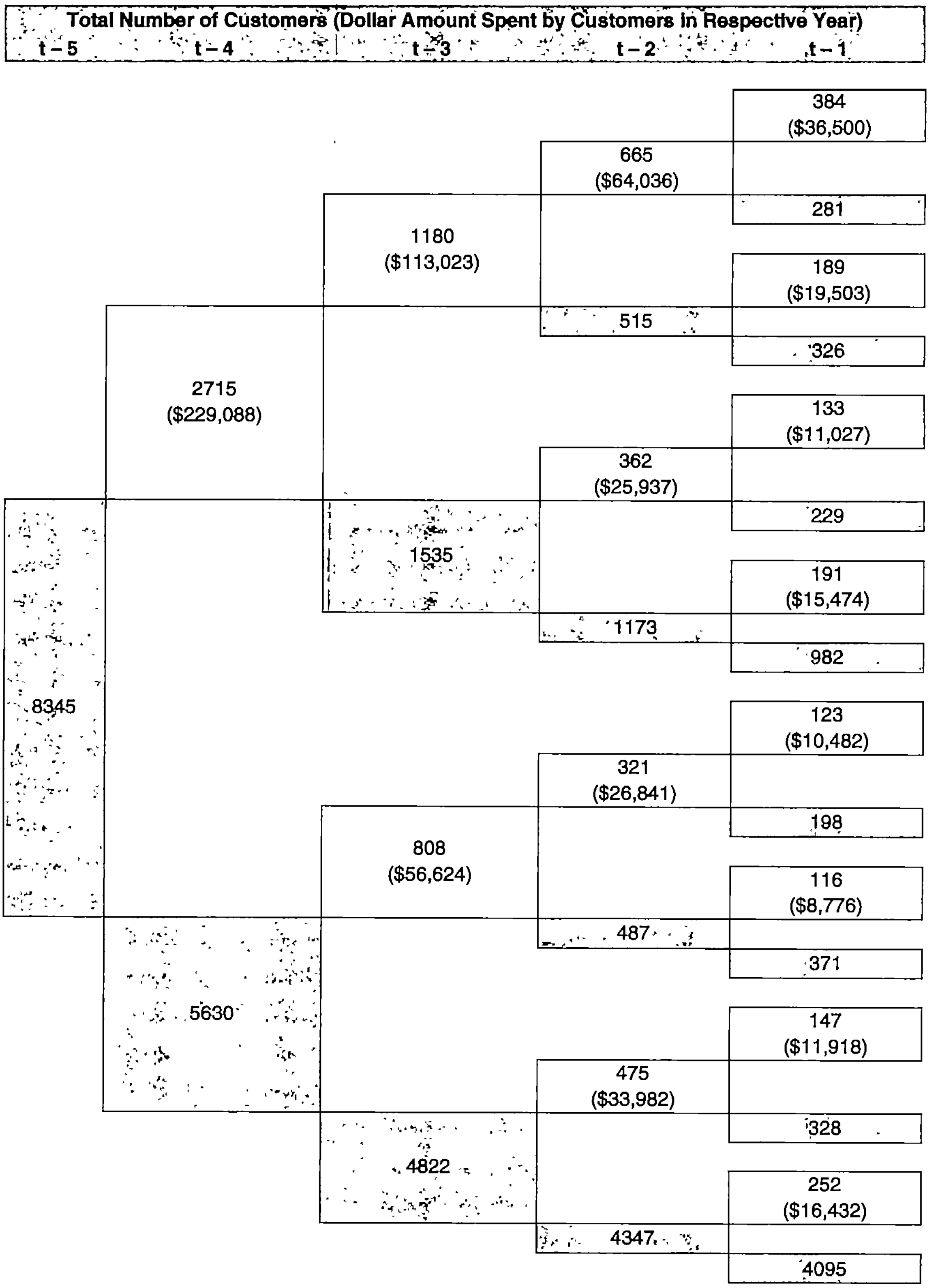
For each of these 36 customer segments, we subsequently calculate conditional purchase frequencies and average dollar amounts spent for all possible purchasing patterns between years t – 4 and t – 1 (for a similar approach, see Berger, Weinberg, and Hanna 2003). Figure 3 visualizes the purchasing pattern of one segment, namely, customers with a recency and frequency of 1 (i.e., customers who last purchased in t – 5 and conducted exactly one purchase between t – 5 and t – 12). For ease of notation, we refer to this segment as "Segment 1" in the following discussion. In this graph, a purchase in one year is represented by an upward move and a white cell, whereas no purchase is represented by a downward move and gray-shaded cell. From the total of 8345 customers in this segment, 2715 conducted a purchase in t – 4, and 5630 did not. Thus, the conditional purchase frequency for a customer to carry out a purchase in t – 4 is 32.53% (2715/8345). Because these 2715 customers spent a total of $229,088, the associated average dollar amount spent is $84.38 per customer ($229,088/2715). In total, there are 15 conditional purchase frequencies and average amounts spent (corresponding to the 15 white cells) per customer segment that can be determined this way. The resulting 15 pairs of values serve as a proxy for future purchasing behavior in t to t + 3 for all consumers who belong to this specific segment.

The second step consists of the estimation of costs generated per customer per year. Specifically, we include two types of costs in our model: the cost of goods sold and direct marketing expenditures. We assume that the cost of goods sold represents a certain percentage of the average dollar amount spent per year, and we estimate this percentage to be approximately 55% on the basis of data available in our data set. For direct marketing expenditures, we face the problem that the data set does not provide sufficient information for their estimation. Therefore, we build on the work of Gönül and Shi (1998, p. 1256), who assume an average cost per mailing of $2.50. On the basis of the assumption that our specialty catalog company might send out between zero and four catalogs per year (i.e., at a maximum of one catalog per season), we assume direct marketing expenditures to vary between $0 and $10 per customer per year. Because there is no way to determine which of these values is more likely than the others, we calculate 100 different scenarios by increasing cost in steps of $.10, beginning at $0 and continuing to $10.[2]

---

[2] We allow marketing expenditures to vary in $.10 increments instead of assuming five distinct values ($.00, $2.50, $5.00, $7.50,

FIGURE 3
Purchasing Pattern of Customer Segment 1 In Years t – 4 to t – 1



In the third step, we use these two inputs to estimate the expected future revenue and profit contribution of an average customer in each segment. We first calculate expected revenue as the average amount spent given a certain purchasing pattern multiplied by the probability that this purchasing pattern can be observed. We then calculate profit contribution as expected revenue less the cost of goods sold less marketing expenditures. Marketing expenditures are weighted with the probability of observing last year's purchasing pattern because we assume that they are paid out at the beginning of each year. Figure 4 shows the resulting profit contributions for an average customer in Segment 1, determined on the basis of the information provided in Figure 3 and the assumption that marketing expenditures are equal to $5 per customer per year.

In the fourth step, we calculate CLV that includes and CLV that excludes option value. We discount profit contributions back to the beginning of year t and sum them up, which results in a CLV of an average customer in Segment N that excludes option value CLV^N_{EOV}. We make the simplifying assumption that profit contributions are realized in the middle of each year because expected revenue can occur at any time, whereas we assume that marketing expenditures are paid out at the beginning of each year. Using a discount rate of 20% and the example of Segment 1, CLV^1_{EOV} is equal to $12.91. To calculate CLV that includes option value CLV^N_{IOV}, we must first determine the activity status of each customer conditional on his or her purchasing pattern. To do this, we make two assumptions: (1) that a customer should be abandoned in year T, whenever the sum of expected future profit contributions from T onward discounted back to T is less than the cost associated with the abandonment, and (2) that an abandoned customer becomes inactive for all years after abandonment. Then, it is straightforward to define a variable that describes the activity level of each customer conditional on his or her purchasing pattern, which takes the value of 1 in the case of the customer being active and 0 in the case of abandonment. In the specific case of Segment 1, this rule results in the decision that a customer who did not conduct a purchase during time t should be abandoned. Thus, CLV^N_{IOV} can be defined as the discounted sum of expected profit contributions multiplied by the corresponding activity level. On the basis of a discount rate of 20% and the assumption that abandonment is not associated with any cost, CLV^1_{IOV} is equal to $13.65 for the example of Segment 1. To obtain an estimate of the divergence between these two values of CLV on customer base level, we multiply CLV^N_{EOV} and CLV^N_{IOV} by the total number of customers in Segment N and sum them over N. Although this is not strictly necessary—a comparison could also be carried out between CLV^N_{EOV} and CLV^N_{IOV} directly—we carry out this calculation to obtain a weighted average of these two values over all segments.

Figure 5 shows the divergence (calculated as $CLV_{IOV}/CLV_{EOV} - 1$) between these two values as a function of marketing expenditures per customer per year. The result supports our hypothesis that CLV that excludes option value underestimates customer value. For marketing expenditures equal to $5 per customer per year, the divergence is ~1.5% on average, and for marketing expenditures greater than $6.50, it consistently exceeds 5%. For any given marketing expenditure, the divergence at the customer segment level differs greatly from one segment to another. For example, in the case of marketing expenditures of $7.50 per customer per year, segment-specific divergence varies from 0% to 70%. This latter point shows that there may be factors other than cost that determine the divergence. Therefore, we conduct a Monte Carlo simulation to investigate this point further.[3]

Monte Carlo Simulation

The purpose of our Monte Carlo simulation is to gain a better understanding of the influence of factors other than cost on the divergence between CLV that includes and CLV that excludes option value. Specifically, we focus on the potential impact of a change in expected revenues, which we modify by changing the standard deviations and average values for the probability of purchase and dollar amount expected to be spent. The simulation consists of two steps: First, we analytically identify which kind of influence these factors can be expected to have on the divergence. Second, we conduct a set of Monte Carlo simulations to determine the order of magnitude of this influence.

For the first step, the analytical identification of the kind of influence to be expected, we build on the equations that Black and Scholes (1973) derive to value financial options. The Black–Scholes model has the advantage of providing a closed-form solution for the value of a (real) option, which is not given either for contingent claims valuation or for dynamic programming. Although we applied dynamic programming, not the Black–Scholes model, as a valuation technique in our empirical validation and the following Monte Carlo simulation, we are comfortable with this choice in the context of the analytical verification. The Black–Scholes model is included as a limiting case in contingent claims valuation (Cox, Ross, and Rubinstein 1979), and contingent claims valuation and dynamic programming are equivalent modulo parameter calibrations (Knudsen, Meister, and Zervos 1999).

Black and Scholes (1973) show (not building on dynamic programming but on the solution of heat-transfer equations in physics) that the value of a put option $V_{RO,P}$ (i.e., the right to sell a stock) can be calculated using the following three equations:[4]

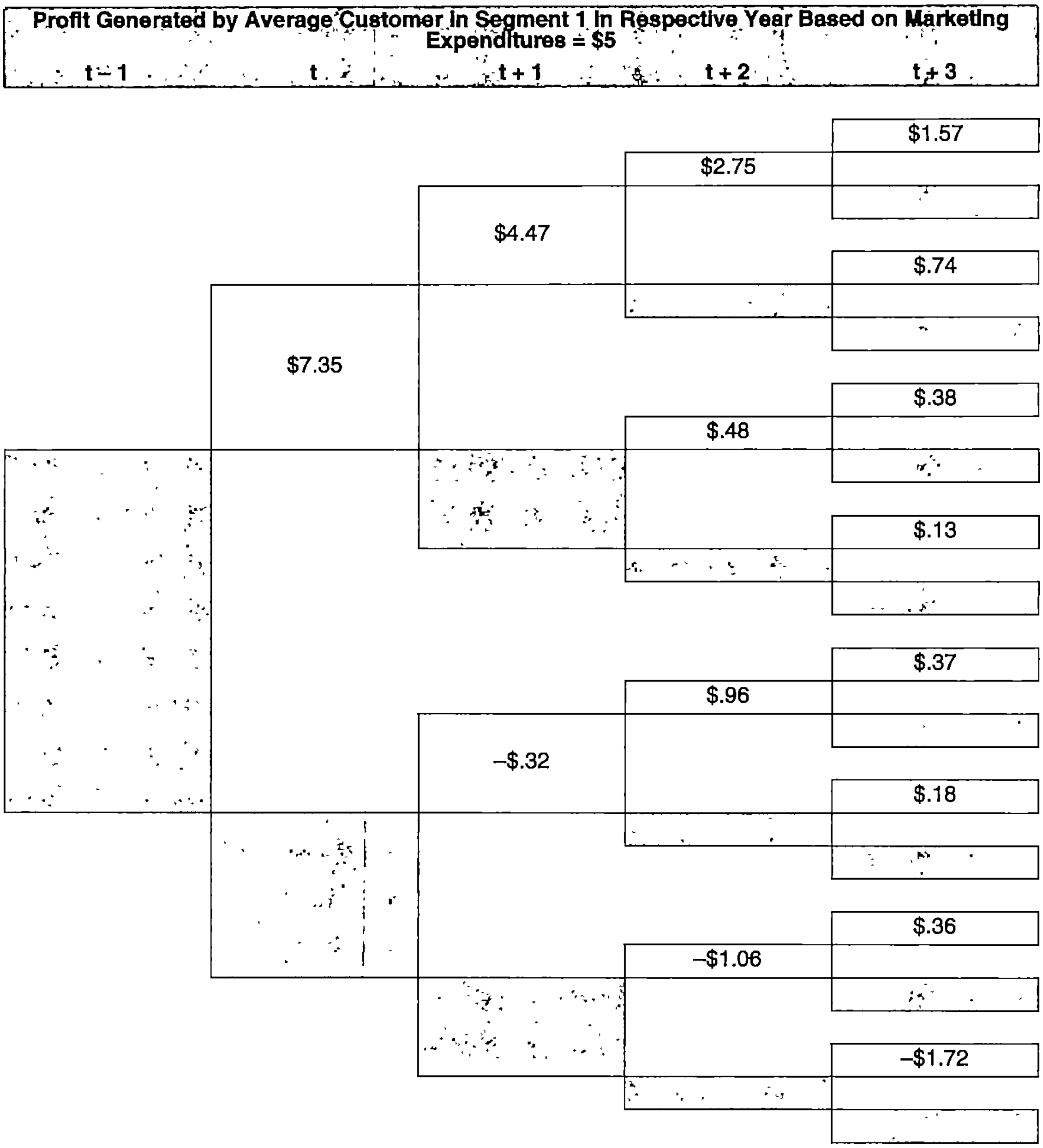
---

and $10.00) to allow for cases in which average cost per mailing differs from $2.50 and/or more than four mailings are sent out per customer per year. We assume that these expenditures are uniform across customers and include both fixed and variable costs. However, they are limited to retention cost only because we assume that no acquisitions take place in the analysis period

[3] We also conduct a similar analysis in which cost of goods sold represent 50% and 60% of revenue. The results show that an increase (decrease) in cost of goods sold simply results in an upward (downward) shift of the graph without substantially influencing variance in segment-specific divergence.

[4] We use the equations that describe a put versus a call option because the real option of abandonment is more similar to a put option in financial terms (Sachdeva and Vandenberg 1993).

FIGURE 4
Profit Contribution of an Average Customer in Segment 1


Profit Generated by Average Customer in Segment 1 in Respective Year Based on Marketing Expenditures = $5
t − 1
t
t + 1
t + 2
t + 3
$1.57
$2.75
$4.47
$.74
$7.35
$.38
$.48
$.13
$.37
$.96
−$.32
$.18
$.36
−$1.06
−$1.72

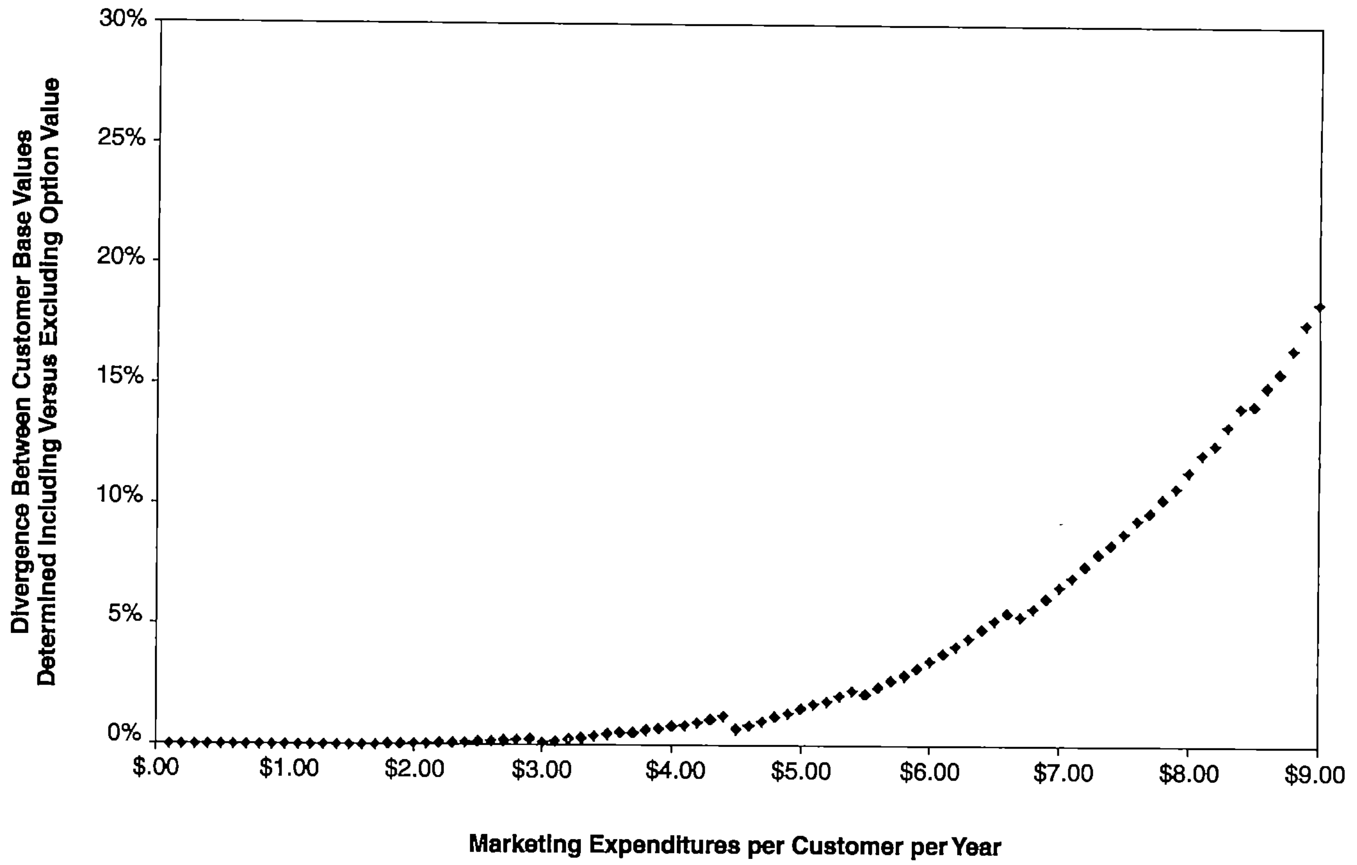
(1) $$V_{RO,P} = -x\,N(-d_1) + ce^{-rt}N(-d_2), \text{ where}$$

(2) $$d_1 = \frac{\ln\frac{x}{c} + \left(r + \frac{1}{2}v^2\right)t}{v\sqrt{t}}, \text{ and}$$

(3) $$d_2 = \frac{\ln\frac{x}{c} + \left(r - \frac{v^2}{2}\right)t}{v\sqrt{t}} = d_1 - v\sqrt{t}.$$

In this expression, x represents the stock price, c represents the exercise price, r represents the risk-free interest rate, v^2 represents the variance rate of the return on the stock, and t represents the time to expiry. The notation ln(a) is used to describe the natural logarithm, and N(a) is used to represent the cumulative normal density function. In the context of real-options analysis, this interpretation changes slightly. As Luehrman (1998) discusses, x can then be viewed as the present value, v^2 can be viewed as the riskiness of expected cash flows, and c can be interpreted as the present value of fixed costs. As we stated previously, the focus of our Monte Carlo simulation is to investigate the potential impact of a change in the standard deviations and average values assumed for the probabilities of purchase and dollar amounts expected to be spent. Because expected revenues can be determined as the product of these two

FIGURE 5
Divergence Between Customer Base Values Determined by Including Versus Excluding Option Value
(Empirical Validation)



parameters, this is equal to investigating a change in the standard deviations and average values of expected revenues. Using the terminology of the Black–Scholes model, we focus on the potential impact of changes in v^2 and x on real-options value.

To derive the impact of these two variables on option value analytically, we calculate the partial derivatives of $V_{RO,P}$ with respect to v and x. Keeping in mind that

$$\frac{\partial N(a)}{\partial a} = f(a) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}a^2}, \tag{4}$$

where f(a) is the normal density function, it follows that

$$\frac{\partial V_{RO,P}}{\partial v} = -\frac{\partial d_1}{\partial v}[ce^{-rt}f(-d_2) - xf(-d_1)] + ce^{-rt}f(-d_2)\sqrt{t} \tag{5}$$

and that

$$\frac{\partial V_{RO,P}}{\partial x} = -\left\{N(-d_1) - \frac{\partial d_1}{\partial x}[xf(-d_1) - ce^{-rt}f(-d_2)]\right\}. \tag{6}$$

Because $d_2 = d_1 - v\sqrt{t}$ and

$$f(-d_2) = f(-d_1)\frac{x}{c}e^{rt}, \tag{7}$$

this can be simplified to

$$\frac{\partial V_{RO,P}}{\partial v} = ce^{-rt}\sqrt{t}f(-d_2) \text{ and} \tag{8}$$

$$\frac{\partial V_{RO,P}}{\partial x} = -N(-d_1). \tag{9}$$

In this simple form, the partial derivative of $V_{RO,P}$ with respect to v has a positive sign, and the one with respect to x has a negative sign. Thus, $V_{RO,P}$ increases with increasing v and decreasing x. In other words, the value of the real option of abandonment increases with increasing uncertainty and decreasing present value of expected cash flows. Translating this back to the factors we investigated in the context of the Monte Carlo simulation, this implies that the difference between CLV_{IOV} and CLV_{EOV} can be expected to increase with increasing standard deviations and decreasing average values of future purchase probabilities and dollar amounts expected to be spent.

In the second step, we then carry out a set of Monte Carlo simulations to get an idea of the order of magnitude of these influences. For this, we build on the always-a-share CLV example that Dwyer (1989) proposes. Dwyer's approach is essentially equivalent to the one we used in the empirical validation; it is characterized by the following two assumptions: First, purchase probabilities, dollar amounts expected to be spent, and marketing expenditures are assumed to depend on the recency cell of each customer (see Table 1). Second, cost of goods sold is assumed to represent 60% of revenue. Dwyer's example is more realistic than our empirical validation in some dimensions (i.e., if we assume that marketing expenditures depend on each customer's recency rather than that they are constant, as in our case) and less in others (i.e., if we assume that purchase probabilities and dollar amounts expected to be spent depend only on recency rather than that purchase probabilities and dollar amounts expected to be spent depend on recency and frequency, as in our analysis). However, because the focus of our Monte Carlo simulation is to investigate how the divergence between CLV that includes and CLV that excludes option value changes with a variation in model parameters, it is not essential that the simulation is completely identical to our empirical validation. It is necessary only to ensure that the assumptions underlying it change in a controlled manner among different scenarios. However, it should be kept in mind that the absolute values of divergence are not comparable. In total, we carried out 1.8 million simulations, split into nine scenarios with 200,000 runs each, to investigate how the divergence changes when we change (1) the standard variation in future purchase probabilities and dollar amounts expected to be spent and (2) the average values assumed for these two parameters.

To address the first objective, the investigation of the change in the divergence associated with a change in standard deviations, we assume that the eight values for purchase probabilities and dollar amounts expected to be spent (one probability and one dollar amount for each of the four recency cells) behave according to a normal distribution with average values as assumed by Dwyer (1989; see Table 1) and standard deviations that represent 5% of these average values. For modeling purchase probabilities, we used a truncated normal distribution to avoid probabilities that were inferior to 0 and larger than 1. We then conducted the following three-step procedure: First, we simulated 20,000 customers by calculating probabilities and dollar amounts expected to be spent on the basis on these distributional assumptions. Second, we determined the resulting divergence between CLV that includes and CLV that excludes option value for each of them. We excluded all customers for which $CLV_{EOV} < 0$ because this resulted in $CLV_{IOV} = 0$ and a divergence of –100%. Third, we calculated an overall average divergence for all 20,000 customers as the median of all customer-level divergences. We decided on the median rather than the arithmetic mean to obtain a conservative estimate and to avoid biases caused by outliers as far as possible. We subsequently increased the standard deviations in 5% increments up to 50% of the average values and followed the same three-step procedure described previously. The results of these 200,000 runs (10 different standard deviations with 20,000 runs reach) appear in Figure 6, in which we plotted the resulting overall average divergence as a function of the standard deviation.

Figure 6 shows two interesting properties: First, there appears to be a quadratic relationship between the average divergence and the standard deviation. This is also confirmed by an associated analysis we conducted using SPSS. Regressing overall average divergence on standard deviation results in an adjusted R-square of .735, an F value of 26.01, and a t-value of the dependent variable of 5.10, whereas regressing it on the squared standard deviation (i.e., variance) results in an adjusted R-square of .909, an F value of 91.35, and a t-value of 9.56. This relationship is intuitive because increasing variance is associated with increasing probabilities for unprofitable customer relationships and, thus, increasing occurrence (and value) of abandonment decisions. Second, the absolute impact of standard deviation on average divergence seems to be marginal. Although standard deviation increases tenfold from 5% to 50%, the average divergence increases only slightly (from .85% to 1.21%). In summary, this analysis confirms our theoretical finding that average divergence increases with increasing standard deviation. Nevertheless, this increase seems to be of limited size.[5]

With respect to the second objective, the analysis of a potential influence of a change in average values on a change in overall average divergence, we again use Dwyer's (1989) assumptions as a starting point. We subsequently modify the average values for purchase probabilities and dollar amounts spent by either doubling or halving the values. This leads to seven other scenarios; for each scenario, we conducted the same analysis described previously (i.e., increasing the standard deviation in 5% increments from 5% to 50% and determining an overall average divergence), the results of which appear in Table 2. The overall average divergence increases with decreasing purchase probabilities and dollar amounts spent, which is in line with our theoretical expectations. For the range of values we investigated in our simulation, this effect seems to be substantially larger than the impact of standard deviation on average diver-

TABLE 1

Assumptions of Always-a-Share CLV Example as Proposed by Dwyer (1989)

Recency Cell	Description	Probability of Purchase in Period t (%)	Amount Expected to Be Spent in Period t ($)	Marketing Expenditures in Period t ($)
1	Customer last purchased in t – 1	30	100	3.60
2	Customer last purchased in t – 2	20	80	3.10
3	Customer last purchased in t – 3	15	60	1.80
4	Customer last purchased in t – 4	5	40	1.30

---

[5] This result was also confirmed in the seven other scenarios we calculated. Because of space constraints, we show only one graph, but this additional information can be obtained from the first author on request.

FIGURE 6
Divergence Between Customer Base Values Determined by Including Versus Excluding Option Value (Simulation, Average Values as Noted in Table 1)

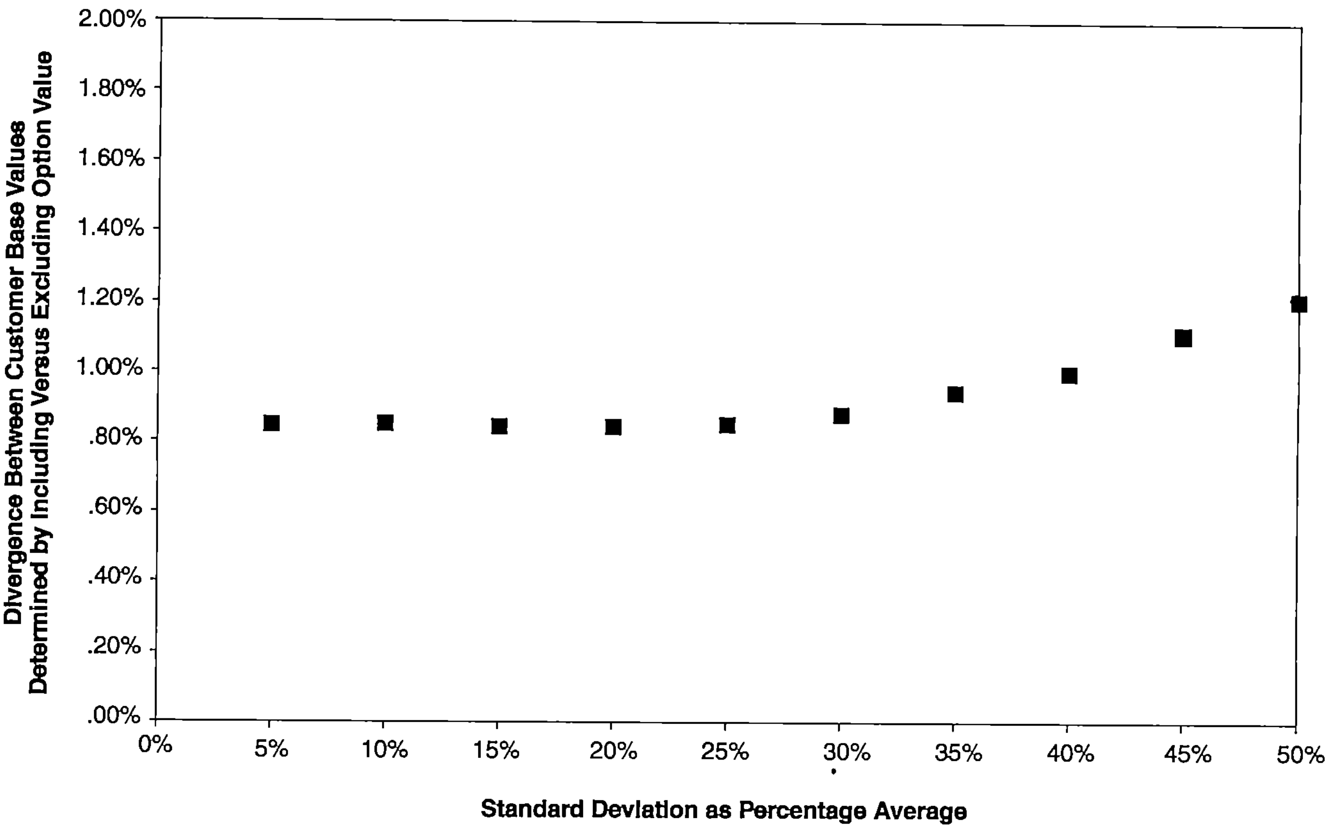


TABLE 2
Divergence Between Customer Base Values Determined by Including Versus Excluding Option Value (Simulation)

	Dollar Amount Expected to Be Spent		
Purchase Probability	**Halved**	**Unchanged**	**Doubled**
Halved		11.52%	1.58%
Unchanged	5.99%	.94%	.03%
Doubled	.25%	.00%	.00%

Notes. "Halved," "unchanged," and "doubled" refer to the average values in Table 1. We did not calculate an overall average divergence for the scenario in which both purchase probabilities and dollar amounts spent were halved, because in this case, for all standard deviations, at least 80% of all simulated customers resulted in $CLV_{EOV} < 0$.

gence. In addition, a change in average purchase probabilities has a greater impact than a change in average dollar amounts spent.

Theoretical Implications and Managerial Relevance

In summary, several authors have used CLV to develop models that focus on the allocation of scarce marketing resources (e.g., Reinartz, Thomas, and Kumar 2005; Rust, Lemon, and Zeithaml 2004; Venkatesan and Kumar 2004). However, it has been discussed in the literature that simple NPV analysis may not reflect the value of the flexibility to make such decisions. Therefore, some authors recommend the use of real-options analysis in certain situations (e.g., Amram and Kulatilaka 1999; Dixit and Pindyck 1995; Pindyck 1991). Building on this stream of research, we proposed a real-options approach for customer valuation. Using the case of the option to abandon unprofitable customers, we first illustrated through example the basic idea behind real-options analysis and its application to determine CLV. We then used data provided by the DMEF to validate our approach empirically and to obtain an estimate of the divergence between CLV that includes and CLV that excludes option value. Finally, we conducted a set of Monte Carlo simulations to obtain a better understanding of factors that influence this divergence. We showed both analytically and through our simulation that the divergence increases with (1) decreasing future purchase probabilities, (2) decreasing dollar amounts expected to be spent in future purchases, and (2) increasing variance in these two parameters. Regarding the order of magnitude of these influences, our simulations provide evidence that the impact of a change in variance is marginal compared with a change in average values. In addition, a change in average purchase probabilities has a larger impact than a change in average dollar amounts spent.

From a theoretical perspective, our approach contributes to research in the area of CLV by proposing a methodology that can be used to incorporate the value of the seller's flexibility into CLV. In this sense, it adds to the literature in three ways: First, it complements Levett and colleagues' (1999) publication and adds to Rust, Lemon, and Zeithaml's (2004) work, both of which explicitly model the customer's flexibility when determining CLV. Second, it addresses two research directions that Hogan and Hibbard (2002) propose; specifically, by using the example of the option to abandon unprofitable customers, we developed a metric to quantify the value of business relationships using the real-options-based framework by explicitly modeling the decision to "fire a customer." Finally, the current study can be viewed in the context of Hogan, Lemon, and Libai's (2003) article because we provide evidence that customers who appear to be unprofitable from an NPV-based CLV perspective may not be as unattractive as they seem.

Therefore, on the basis of our findings, we recommend including option value in the calculation of CLV. Our results show that neglecting option value can lead to a substantial underestimation of true customer value. Because this underestimation is not uniform across the whole customer base, it is likely to affect decisions made on the basis of a rank order of customers developed using CLV. The results of our Monte Carlo simulation show that customers with low probabilities of future purchase and low dollar amount expected to be spent tend to be underestimated more often than customers with more favorable values along these two dimensions. Thus, a rank order of customers based on CLV may be different when CLV is calculated including option value than when it is calculated excluding option value. Most methods for the allocation of scarce marketing resources recommended in the literature (e.g., Reinartz, Thomas, and Kumar 2005; Rust, Lemon, and Zeithaml 2004; Venkatesan and Kumar 2004) build on such a rank order of customers and recommend the allocation of more resources to customers with higher ranks. However, our results lead to the hypothesis that this allocation could be substantially different if option value were included.

With respect to the implementability of our approach, we showed in our example and our empirical validation that the calculation of CLV that includes option value does not require additional data compared with the determination of CLV that excludes option value. Many companies already classify customers into a set of distinct RFM cells. Thus, the basic information to determine option value is most likely already available in any given company. If it is not available, it can easily be generated if information about the purchase history of each individual customer is known. Building on this allocation of customers into distinct RFM cells, it is possible to determine option value using standard spreadsheet software. We conducted all calculations in this article using Excel 2002.

Regarding the practicability of abandoning unprofitable customers, approaches have been discussed in the literature for a long time under the label "selective demarketing" (Kotler and Levy 1971). With the increasing importance of CLV, the topic has recently received the attention of a wider range of researchers (e.g., Rosenblum, Tomlinson, and Scott 2003; Selden and Colvin 2003; Zeithaml, Rust, and Lemon 2001). Techniques for dealing with unprofitable customers that have been proposed in this context can be grouped into "hard" and "soft" approaches. Hard approaches imply telling the customer directly that he or she is abandoned. For example, a recent article in *The Seattle Times* (Freed 2004) speaks about the Massachusetts-based fashion retailer Filene's Basement. This company banned two sisters, who claimed to have been loyal customers for years, from all 21 of its stores after deciding that they returned too many items after wearing them for an evening and complained too often about service. In contrast to this strategy, soft approaches summarize all actions that can be taken to make the customer leave on his or her own. Among others, this includes not sending any promotional materials to such customers, introducing additional fees (e.g., chargeable customer service calls), and stopping certain services (Selden and Colvin 2003).

Limitations and Areas of Further Research

Our study has two limitations that further research could address: First, our findings apply only to the specific type of option we analyzed, namely, the abandonment option. Thus, the degree of divergence we determined during our empirical validation and the factors that influence it should not be generalized to other types of real options. We chose this specific option type as an illustration because some authors argue that the flexibility that lies at the core of the real-options approach requires the possibility of abandoning investment initiatives. This perspective makes abandonment essential for limiting downside risk (Adner and Levinthal 2004). It also differentiates abandonment options from other option types that have been discussed in the literature (i.e., defer, learning, growth, and switching options). However, our approach can be modified easily to reflect these types of options. Nevertheless, further analysis is needed to obtain a better understanding of this type and other types of options.

Second, our empirical analysis and Monte Carlo simulation both assume that abandoning a customer is not associated with any cost. However, this is unlikely to be the case. Although it may be possible to cease business with an unprofitable customer without creating additional direct cost, abandonment carries the risk of creating an image of unfairness and discrimination in the market, which could lead to substantial indirect cost (Cullwick 1975). For example, previous research has shown that companies that behave unfairly may be punished in the long run, even when this results in cost to their customers (Kahneman, Knetsch, and Thaler 1986). One such punishment could be the spread of negative word of mouth, the effects of which can be substantial if the number of customers who experience dissatisfaction is high enough (Richins 1983). Such indirect abandonment costs should be considered, perhaps by an approach similar to the one that Hogan, Lemon, and Libai (2004) use to quantify positive word of mouth.

In addition to these two areas of further research, we believe that studies calculating real-options value by making use of more sophisticated approaches to model future customer behavior, such as the Pareto/NBD (Schmittlein, Morrison, and Colombo 1987) or the beta-geometric/NBD (Fader, Hardie, and Lee 2005a) framework, would be worthwhile. Such methods could help account for heterogeneity among different customers more effectively than the RFM approach we used in our empirical validation, which could contribute to a better understanding of customer-specific factors that influence option value. Furthermore, this line of research would be an initial step in marrying our approach with the work done regarding the calculation of CLV based on these approaches (e.g., Fader, Hardie, and Lee 2005b; Reinartz and Kumar 2000, 2003). Finally, we believe that combining our work with publications focused on the integration of the buyer's flexibility into CLV (e.g., Levett et al. 1999; Rust, Lemon, and Zeithaml 2004) to come up with an integrated model that simultaneously values the buyer's and seller's options would be a promising area of further research. Such a model could be helpful in two ways: First, it could further improve the understanding of the "true value of a lost customer" (Hogan, Lemon, and Libai 2003) and, consequently, the quality of decisions regarding customer selectivity by better distinguishing between good and bad customers. Second, it could contribute to firms making better recommendations about the allocation of scarce marketing resources based on a more profound understanding of "return on marketing" (Rust, Lemon, and Zeithaml 2004).

Appendix

In the following notation, let X^N_{ijkl} describe a variable, X, which applies to an average customer of Segment N (N = 1, 2, …, 36) with purchasing pattern ijkl. Here, i(j, k, l) takes the value of 1 in the case of a purchase having taken place in year t (t + 1, t + 2, t + 3) and 0 otherwise. Furthermore, let COGS represent cost of goods sold as a percentage of sales and ME stand for fixed marketing expenditures per customer per year that are assumed to be paid at the beginning of each year. Finally, let f^N_{ijkl} and AA^N_{ijkl} describe the purchase frequencies and average amounts spent determined on the basis of the RFM approach.

Using these notations, we can express the expected revenue, ER^N_{ijkl}, and profit contribution, PC^N_{ijkl}, as follows:

$$\text{(A1)}\quad PC^N_i = ER^N_i \times (1 - COGS) - ME, \text{ with}$$

$$ER^N_i = f^N_i \times AA^N_i E;$$

$$\text{(A2)}\quad PC^N_{il} = ER^N_{il} \times (1 - COGS) - f^N_i \times ME, \text{ with}$$

$$ER^N_{il} = f^N_i \times f^N_{il} \times AA^N_{il};$$

$$\text{(A3)}\quad PC^N_{ijl} = ER^N_{ijl} \times (1 - COGS) - f^N_i \times f^N_{ij} \times ME, \text{ with}$$

$$ER^N_{ijl} = f^N_i \times f^N_{ij} \times f^N_{ijl} \times AA^N_{ijl}; \text{ and}$$

$$\text{(A4)}\quad PC^N_{ijkl} = ER^N_{ijkl} \times (1 - COGS) - f^N_i \times f^N_{ij} \times f^N_{ijk} \times ME, \text{ with}$$

$$ER^N_{ijkl} = f^N_i \times f^N_{ij} \times f^N_{ijk} \times f^N_{ijkl} \times AA^N_{ijkl}.$$

On the basis of Equations A1–A4, it is now straightforward to calculate CLV that excludes option value. If we take a discount rate of d% per year and, for the sake of simplicity, assume that the profit contribution of each customer is paid out in the middle of each year (expected revenue can occur at any time, whereas marketing expenditures are assumed to be paid out at the beginning of each year), we can calculate the CLV for an average customer in Segment N based on a lifetime of four years (t to t + 3), $CLV^N{}_{EOV}$ as follows:

$$\text{(A5)}\quad CLV^N_{EOV} = \frac{\sum_{ijk} PC^N_{ijkl}}{(1+d)^{3.5}} + \frac{\sum_{ij} PC^N_{ijl}}{(1+d)^{2.5}} + \frac{\sum_{i} PC^N_{il}}{(1+d)^{1.5}} + \frac{PC^N_l}{(1+d)^{.5}}.$$

To determine CLV that includes option value, it is first necessary to define a decision rule that provides information about when a customer should be abandoned. Let $DV^{N,T}{}_{ijkl}$ be the value of an average customer of Segment N at the beginning of year T (T = t + 1, t + 2, t + 3), which we can calculate as the sum of expected future profit contributions from T onward, discounted back to T, according to the following three formulas:

$$\text{(A6)}\quad DV^{N,t+1}_i = \frac{\sum_{jk} PC^N_{ijkl}}{(1+d)^{2.5}} + \frac{\sum_{j} PC^N_{ijl}}{(1+d)^{1.5}} + \frac{PC^N_{il}}{(1+d)^{.5}},$$

$$\text{(A7)}\quad DV^{N,t+2}_{ij} = \frac{\sum_{k} PC^N_{ijkl}}{(1+d)^{1.5}} + \frac{PC^N_{ijl}}{(1+d)^{.5}}, \text{ and}$$

$$\text{(A8)}\quad DV^{N,t+3}_{ijk} = \frac{PC^N_{ijkl}}{(1+d)^{.5}}.$$

A customer should be abandoned at the beginning of year T whenever $DV^{N,T}{}_{ijkl}$ is less than the cost associated with this abandonment AC. Let $AB^{N,T}{}_{ijkl}$ be a variable that takes the value of 1 when the customer should be abandoned at the beginning of year T and 0 otherwise. Using Equations A5–A8, we can determine $AB^{N,T}{}_{ijkl}$ as follows:

(A9) $$AB^{N,t} = \begin{cases} 1 & \text{if } CLV^{N}_{NPV} < AC \\ 0 & \text{if otherwise} \end{cases},$$

(A10) $$AB^{N,t+1}_{i} = \begin{cases} 1 & \text{if } DV^{N,t+1}_{i} < AC \\ 0 & \text{if otherwise} \end{cases},$$

(A11) $$AB^{N,t+2}_{ij} = \begin{cases} 1 & \text{if } DV^{N,t+2}_{ij} < AC \\ 0 & \text{if otherwise} \end{cases}, \text{ and}$$

(A12) $$AB^{N,t+3}_{ijk} = \begin{cases} 1 & \text{if } DV^{N,t+3}_{ijk} < AC \\ 0 & \text{if otherwise} \end{cases}.$$

Taking into account that a customer, after he or she is abandoned, turns from an active to an inactive status for all future years, we can define a variable, $S^{N,T}{}_{ijkl}$, which defines the activity status of a customer at year T and takes either the value of 1 when the customer is active or the value of 0 when the customer is inactive, as follows:

(A13) $$S^{N,t} = 1 - AB^{N,t},$$

(A14) $$S^{N,t+1}_{i} = 1 - \min(S^{N,t} + AB^{N,t+1}_{i};\ 1),$$

(A15) $$S^{N,t+2}_{ij} = 1 - \min(S^{N,t+1}_{i} + AB^{N,t+2}_{ij};\ 1), \text{ and}$$

(A16) $$S^{N,t+3}_{ijk} = 1 - \min(S^{N,t+2}_{ij} + AB^{N,t+3}_{ijk};\ 1).$$

Finally, building on Equations A5 and A13–A16, we can now easily calculate CLV that includes option value as follows:

(A17) $$CLV^{N}_{IOV} = \frac{\sum_{ijk} S^{N,t+3}_{ijk} PC^{N}_{ijk1}}{(1+d)^{3\ 5}} + \frac{\sum_{ij} S^{N,t+2}_{ij} PC^{N}_{ij1}}{(1+d)^{2\ 5}} + \frac{\sum_{i} S^{N,t}_{i} PC^{N}_{i1}}{(1+d)^{1\ 5}} + \frac{S^{N,t} PC^{N}_{1}}{(1+d)^{5}}.$$

REFERENCES

Adler, Ralph W. (2000), "Strategic Investment Decision Appraisal Techniques: The Old and the New," *Business Horizons*, 43 (6), 15–22.

Adner, Ron and Daniel A. Levinthal (2004), "What Is Not a Real Option: Considering Boundaries for the Application of Real Options to Business Strategy," *Academy of Management Review*, 29 (1), 74–85.

Amram, Martha and Nalin Kulatilaka (1999), *Real Options: Managing Strategic Investment in an Uncertain World.* Boston: Harvard Business School Press.

Bauer, Connie L. (1988), "A Direct Mail Customer Purchase Model," *Journal of Direct Marketing*, 2 (3), 16–24.

Benninga, Simon and Efrat Tolkowsky (2002), "Real Options: An Introduction and an Application to R&D Valuation," *The Engineering Economist*, 47 (2), 151–68.

Berger, Paul D., Bruce Weinberg, and Richard C. Hanna (2003), "Customer Lifetime Value Determination and Strategic Implications for a Cruise-Ship Company," *Database Marketing and Customer Strategy Management*, 11 (1), 40–52.

Bitran, Gabriel R. and Susana V. Mondschein (1996), "Mailing Decisions in the Catalog Sales Industry," *Management Science*, 42 (9), 1364–81.

Black, Fischer and Myron Scholes (1973), "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81 (3), 637–54.

Copeland, Tom and Peter Tufano (2004), "A Real-World Way to Manage Real Options," *Harvard Business Review*, 82 (3), 90–99.

Cox, John C., Stephen A. Ross, and Mark Rubinstein (1979), "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 7 (3), 229–63.

Cullinan, George J. (1978), "'Picking Them by Their Batting Averages'": Recency-Frequency-Monetary Method of Controlling Circulation," manual release. New York: Direct Mail/Marketing Association.

Cullwick, David (1975), "Positioning Demarketing Strategy," *Journal of Marketing*, 39 (April), 51–57.

Devinney, Timothy M. and David W. Stewart (1988), "Rethinking the Product Portfolio: A Generalized Investment Model," *Management Science*, 34 (9), 1080–1095.

Dixit, Avinash K. and Robert S. Pindyck (1995), "The Options Approach to Capital Investment," *Harvard Business Review*, 73 (3), 105–115.

Dwyer, F Robert (1989), "Customer Lifetime Valuation to Support Marketing Decision Making," *Journal of Direct Marketing*, 3 (4), 8–15.

Elton, Edwin J. and Martin J. Gruber (1971), "Dynamic Programming Applications in Finance," *Journal of Finance*, 26 (2), 473–506.

Fader, Peter S., Bruce G.S. Hardie, and Ka Lok Lee (2005a), "Counting Your Customers the Easy Way: An Alternative to the Pareto/NBD Model," *Marketing Science*, 24 (2), 275–84.

———, ———, and ——— (2005b), "RFM and CLV: Using Iso-Value Curves for Customer Base Analysis," *Journal of Marketing Research*, 42 (November), 415–30.

Finch, J. Howard, Richard C. Becherer, and Richard Casavant (1998), "An Option-Based Approach for Pricing Perishable Services Assets," *Journal of Services Marketing*, 12 (6), 473–83.

Freed, Joshua (2004), "Customers from Hell Face Retail Retribution," *The Seattle Times*, (July 6), (accessed February 22, 2006), [available at http://archives.seattletimes.nwsource.com/web/index.html].

Gönül, Füsun and Meng Ze Shi (1998), "Optimal Mailing of Catalogs: A New Methodology Using Estimable Structural Dynamic Programming Models," *Management Science*, 44 (9), 1249–62.

Gupta, Sunil, Donald R. Lehmann, and Jennifer Ames Stuart (2004), "Valuing Customers," *Journal of Marketing Research*, 41 (February), 7–18.

Herath, Hemantha S.B. and Chan S. Park (2001), "Real Options Valuation and Its Relationship to Bayesian Decision-Making Methods," *The Engineering Economist*, 46 (1), 1–32.

Hogan, John E. and Jonathan D. Hibbard (2002), "A Real Options-Based Framework for Valuing Business Relationships as Strategic Assets," working paper, School of Management, Boston University.

———, Katherine N. Lemon, and Barak Libai (2003), "What Is the True Value of a Lost Customer?" *Journal of Service Research*, 5 (3), 196–208.

———, ———, and ——— (2004), "Quantifying the Ripple: Word-of-Mouth and Advertising Effectiveness," *Journal of Advertising Research*, 44 (3), 271–80.

Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler (1986), "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *The American Economic Review*, 76 (4), 728–41.

Knudsen, Thomas S., Bernhard Meister, and Mihail Zervos (1999), "On the Relationship of the Dynamic Programming Approach and the Contingent Claim Approach to Asset Valuation," *Finance and Stochastics*, 3 (4), 433–49.

Kotler, Philip (1974), "Marketing During Periods of Shortage," *Journal of Marketing*, 38 (July), 20–29.

——— and Sidney J. Levy (1971), "Demarketing, Yes, Demarketing," *Harvard Business Review*, 49 (6), 74–80.

Levett, Peter, Michael Page, Deon Nel, Leyland Pitt, Pierre Berthon, and Arthur Money (1999), "Towards an Application of Option Pricing Theory in the Valuation of Customer Relationships," *Journal of Strategic Marketing*, 7 (4), 275–84.

Luehrman, Timothy A. (1998), "Investment Opportunities as Real Options: Getting Started on the Numbers," *Harvard Business Review*, 76 (4), 51–67.

Miller, Luke T. and Chan S. Park (2005), "A Learning Real Options Framework with Applications to Process Design and Capacity Planning," *Production and Operations Management*, 14 (1), 5–20.

O'Sullivan, Orla (1997), "Some of Your Customers Are Unprofitable. OK, Now What?" *ABA Banking Journal*, 89 (11), 42–46, 93.

Pindyck, Robert S. (1991), "Irreversibility, Uncertainty, and Investment," *Journal of Economic Literature*, 29 (3), 1110–48.

Reichheld, Frederick F. and W. Earl Sasser Jr. (1990), "Zero Defections: Quality Comes to Service," *Harvard Business Review*, 68 (5), 105–111.

Reinartz, Werner J. and V. Kumar (2000), "On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing," *Journal of Marketing*, 64 (October), 17–35.

——— and ——— (2003), "The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration," *Journal of Marketing*, 67 (January), 77–99.

———, Jacquelyn S. Thomas, and V. Kumar (2005), "Balancing Acquisition and Retention Resources to Maximize Customer Profitability," *Journal of Marketing*, 69 (January), 63–79.

Richins, Marsha L. (1983), "Negative Word-of-Mouth by Dissatisfied Customers: A Pilot Study," *Journal of Marketing*, 47 (January), 68–78.

Rosenblum, David, Doug Tomlinson, and Larry Scott (2003), "Bottom-Feeding for Blockbuster Businesses," *Harvard Business Review*, 81 (3), 52–59.

Rust, Roland T, Katherine N. Lemon, and Valarie A. Zeithaml (2004), "Return on Marketing: Using Customer Equity to Focus Marketing Strategy," *Journal of Marketing*, 68 (January), 109–127.

Sachdeva, Kanwal and Pieter A. Vandenberg (1993), "Valuing the Abandonment Option in Capital Budgeting: An Option Pricing Approach," *Financial Practice and Education*, 3 (2), 57–65.

Schmittlein, David C., Donald G. Morrison, and Richard Colombo (1987), "Counting Your Customers: Who Are They and What Will They Do Next?" *Management Science*, 33 (1), 1–24.

Selden, Larry and Geoffrey Colvin (2003), "M&A Needn't Be a Loser's Game," *Harvard Business Review*, 81 (6), 70–79.

Slater, Stanley F., Venkateshwar K. Reddy, and Thomas J. Zwirlein (1998), "Evaluating Strategic Investments: Complementing Discounted Cash Flow Analysis with Options Analysis," *Industrial Marketing Management*, 27 (5), 447–58.

Venkatesan, Rajkumar and V. Kumar (2004), "A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy," *Journal of Marketing*, 68 (October), 106–125.

Zeithaml, Valarie A., Roland T. Rust, and Katherine N. Lemon (2001), "The Customer Pyramid: Creating and Serving Profitable Customers," *California Management Review*, 43 (4), 118–42.

Christian Homburg, Nicole Koschate, & Wayne D. Hoyer

The Role of Cognition and Affect in the Formation of Customer Satisfaction: A Dynamic Perspective

Despite the strong recognition that customer satisfaction should be viewed from a dynamic perspective, little is known about how the satisfaction judgment develops over time. Therefore, this study provides a dynamic analysis of the simultaneous influence of cognition and affect in the satisfaction formation process. The results of an experimental study based on a real consumption experience indicate that the impact of cognition on the satisfaction evaluation increases and the influence of affect decreases over time. Moreover, these effects are attenuated with inconsistent performance experiences. Finally, the study shows that the variance in customer satisfaction jointly explained by cognition and affect increases as experience accumulates.

The topic of customer satisfaction/dissatisfaction has been of great interest to marketing and consumer researchers for many years. This interest is driven in part by the notion that customer satisfaction can have long-term benefits, including customer loyalty and increased profitability (e.g., Anderson, Fornell, and Lehmann 1994; Anderson, Fornell, and Rust 1997; Anderson and Sullivan 1993; Bearden and Teel 1983; Bolton and Drew 1991; Fornell 1992; LaBarbera and Mazursky 1983; Oliver 1980; Oliver and Swan 1989a, b; Rust, Moorman, and Dickson 2002; Rust and Zahorik 1993). Against this background, it is important to understand the mechanism that leads to customer satisfaction from both an academic and an applied perspective.

Previous research has recognized that both cognition and affect significantly predict satisfaction judgments. Cognition has been studied mainly in terms of the disconfirmation paradigm, which predicts satisfaction to be a function of a comparison between expectations and performance (e.g., Bearden and Teel 1983; LaBarbera and Mazursky 1983; Oliver 1980; Oliver and DeSarbo 1988). Other studies have recognized that the affect experienced during the acquisition and consumption of the product or service (e.g., joy, happiness, disgust) can also have a significant influence on satisfaction judgments (e.g., Mano and Oliver 1993; Westbrook 1987; Westbrook and Oliver 1991). Oliver (1997, p. 319) states that affect "coexists alongside various cognitive judgments in producing satisfaction" and that it is central to understanding customers' consumption experiences. In addition, a key conclusion of Szymanski and Henard's (2001) meta-analysis is that disconfirmation and, to a lesser extent, affect are strongly related to satisfaction.

However, only a few studies have investigated cognitive and affective antecedents of customer satisfaction simultaneously. For example, in studying affects that are provoked by service failures, Smith and Bolton (2002) find that feelings predict satisfaction levels, after accounting for cognitive factors. Kempf's (1999) study finds that feelings (e.g., arousal, pleasure) are particularly important antecedents of product evaluations of hedonic products, whereas brand cognitions are not. The opposite is true for functional products. Oliver (1993) examines the role of disconfirmation and affective experiences and finds that both are critical to the modeling of satisfaction judgments.

Although previous studies identify the key roles of cognition and affect in influencing satisfaction, it is important to acknowledge that previous work has largely been static (i.e., cross-sectional) in nature. This represents a significant research gap because it is well established that customer satisfaction is a dynamic phenomenon. Even the classic confirmation/disconfirmation paradigm, which has guided much of the research in the area, recognizes that satisfaction judgments can have an important influence on customer expectations for subsequent consumption occasions (Oliver 1980). The dynamic nature of customer satisfaction is also emphasized in the work by Bolton (1998), who proposes a dynamic model of cumulative satisfaction. In a similar vein, Bolton and Lemon (1999) stress the importance of a dynamic perspective with their model of customer service usage. Despite this strong recognition that customer satisfaction should be viewed from a dynamic perspective, the role of cognitive and affective influences has not been systematically studied in this manner.

The few studies that investigate the antecedents of customer satisfaction from a dynamic perspective focus on the cognitive component of customer satisfaction. For example,

---

Christian Homburg is Professor of Business Administration and Marketing and Chairman of the Department of Marketing (e-mail: homburg@bwl.uni-mannheim.de), and Nicole Koschate is Assistant Professor of Marketing (e-mail: nicole.koschate@bwl.uni-mannheim.de), University of Mannheim, Germany. Wayne D. Hoyer is James L. Bayless/William S. Farish Fund Chair for Free Enterprise and Chairman of the Department of Marketing, McCombs School of Business, University of Texas at Austin (e-mail: wayne.hoyer@mccombs.utexas.edu) The authors thank the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 504, for financial support.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Mittal, Kumar, and Tsiros (1999) and Mittal, Katrichis, and Kumar (2001) find that the weights given to an attribute in determining satisfaction shift over time on the basis of their salience. Slotegraaf and Inman (2004) show how satisfaction with product attributes can decline over time (particularly those that can be remedied and especially as customers approach the end of their warranty period). Other dynamic studies in this area focus on carryover effects of previous satisfaction judgments to subsequent satisfaction evaluations (Bolton and Drew 1991; Boulding et al. 1993; Mittal, Kumar, and Tsiros 1999). Bolton (1998) proposes a dynamic model in which the duration of customer relationship depends on an anchoring and adjustment process in which cumulative satisfaction serves as the anchor for subsequent experiences. Nevertheless, to the best of our knowledge, there is no study that investigates the effects of cognition and affect simultaneously over time.

Against this background, our study investigates cognition and affect simultaneously as drivers of customer satisfaction in a dynamic context. The fundamental proposition of our study is that the role of cognition and affect may change over time. More specifically, we argue that affect plays its strongest role at early stages of satisfaction development because of a lack of knowledge about the product or service. As experience accumulates, however, the impact of cognition should increase. To study this proposition, we examine customer satisfaction formation, beginning from a point at which customers have no prior experience with the product at all. In this respect, our study differs from previous work in this area, which essentially considers situations in which customers have some prior experience with the product or service. In addition, we examine this process in the context of moderately complex, utilitarian products.

In addition, from a dynamic perspective, the evaluation process that occurs during consumption serves as an input into subsequent decision-making situations. In other words, the nature of the feedback received has a strong influence on the satisfaction judgment, and this affects future decisions. Therefore, we build a case that the phenomenon under study also depends on the level of consistency of the consumption experience (i.e., Is it consistently positive, consistently negative, or inconsistent?). The importance of the concept of feedback consistency has been established in related areas, such as job performance (Alvero, Bucklin, and Austin 2001; Stone and Stone 1985). To the best of our knowledge, however, this concept has not been studied in the customer satisfaction literature.

Furthermore, previous research in social psychology and marketing has indicated that attitudes and judgments become more stable over time after multiple experiences with the attitude object (Fazio 1986; Petty, Unnava, and Strathman 1991; Smith and Swinyard 1983). On the basis of this literature, we argue that over time and with repeated consumption experiences, satisfaction judgments should also become more stable and that the ability of both cognitive and affective factors to predict customer satisfaction should increase.

Thus, our research is in line with the frequent call for longitudinal studies in marketing (e.g., Bolton 1998; Fournier and Mick 1999; Szymanski and Henard 2001). In addition to the theoretical interest of this topic, it has practical relevance because it aids managers in understanding the development of customer satisfaction and what types of factors are most critical at different stages in the process.

Development of Hypotheses

According to Oliver (1997), a satisfaction evaluation is a specific type of attitude (i.e., postconsumption evaluation). Therefore, to develop our hypotheses on the customer satisfaction formation process, we draw on research on attitudes. Research in this area has found that newly formed attitudes are unstable and held with less certainty and therefore are more difficult to predict (Eagly and Chaiken 1993). Certainty has been shown to increase as the quantity of information about the product increases (Berscheid et al. 1976; Dover and Olson 1977; Farley, Katz, and Lehmann 1978). Over repeated consumption experiences, customers gain more information about the product, and therefore their judgment certainty should increase (Fazio 1986; Smith and Swinyard 1983). Thus, with repeated experience, the cognitive and affective judgments underlying these attitudes become associated with higher certainty (i.e., become stronger and more stable) and should be better able to predict the satisfaction evaluation (Chandrashekaran et al. 2000). We hypothesize the following:

> H_1: The variance in customer satisfaction jointly explained by cognitive and affective factors increases as experience accumulates.

As we mentioned previously, a key contribution of this article is its examination of whether the impact of cognition and affect on the satisfaction judgment changes over time. In particular, we are interested in the early stages of the satisfaction formation process (i.e., the first consumption experiences with the product or service). Whenever customers encounter a new product or service, there is typically little information to draw on from memory (i.e., no affect- or cognition-based evaluations). Therefore, customers are more likely to rely on their feelings because feelings are often elicited immediately on exposure to a new stimulus (Pham et al. 2001). In other words, affective responses to stimuli are evoked much quicker than cognitive responses. Consistent with this view, several studies have examined a feelings-based inference called the "how-do-I-feel-about-it" heuristic, whereby customers can make satisfaction or dissatisfaction judgments based on the valence of their feelings (Gorn, Goldberg, and Basu 1993; Pham 1998; Schwartz and Clore 1988).

This notion is also related to strategies outlined in the affect infusion model (AIM; Forgas 1994, 1995), in which affect infusion is highly likely to occur. Affect infusion refers to the process in which affectively loaded information influences and becomes part of the judgmental process, entering the judge's constructive thought process and eventually coloring the judgmental outcome (Forgas 1995, p. 39). The AIM is a cognitively oriented approach for describing the role of affect in judgment processes and thus fits well with the decision context of the current study. The AIM suggests that it is the constructive, generative nature of most judgments that is critical for affect infusion to occur (Fiedler 1991). Constructive processing involves the development of a judgment or decision rule on the spot rather than the reliance on a well-formed rule or judgment (Bettman, Luce, and Payne 1998).

Applied to the current context, for the first consumption experience, customers should have little information stored in their memory on which to base a judgment, and therefore they should rely on affective inputs. As the number of consumption experiences increases, however, customers acquire increasing amounts of information about the product or service (i.e., through trial and other sources). When making a satisfaction judgment on subsequent occasions, customers can recall this prior information to help them make a satisfaction judgment. In support of this notion, studies have found that previous satisfaction judgments affect subsequent satisfaction evaluations (Bolton 1998; Bolton and Drew 1991; Boulding et al. 1993; Mittal, Kumar, and Tsiros 1999). This notion is consistent with the direct access strategy of the AIM model, which involves the strongly cued retrieval of a crystallized judgment from memory, and is a robust process more resistant to affect infusion.

Thus, we expect that affect has a greater impact on the satisfaction judgment in the earlier stages of the satisfaction development process when judgments require a higher degree of constructive processing. In support of this notion, Smith and Bolton (2002) find that affect is more likely to predict transaction-specific customer satisfaction than cumulative satisfaction (which would be based on several consumption experiences). Thus, we offer the following hypothesis:

> H_2: As experience accumulates, (a) the impact of affective factors on customer satisfaction decreases, and (b) the impact of cognitive factors on customer satisfaction increases.

Further refining our theoretical reasoning, we propose that the previously mentioned pattern of processing is contingent on the consistency of the consumption experience over time. Thus, we investigate the impact of consistent versus inconsistent performance experiences on the proposed relationships in H_2. A consistent experience would result from performance feedback that is in the same direction for each consumption occasion (i.e., all consumption experiences are positive, or all are negative). This leads to the building of a more reliable and diagnostic knowledge base that can be used as a basis for evaluation on subsequent occasions. In support of this notion, studies have found that as consistency of information increases, attitude certainty also increases (Heslin, Blake, and Rotton 1972; Kahneman and Tversky 1973). Therefore, we expect that customers switch in early stages from a more constructive processing strategy to a direct access strategy.

Conversely, inconsistent experiences would generate conflicting information, which would result in less reliable and less diagnostic information stored in memory. Therefore, customers would need to continue engaging in constructive information processing, which, as we mentioned previously, involves the significant role of affective factors. Thus, we assume that the relative impact of affective factors does not decrease as heavily as in the case of consistent experiences.

> H_3: The suggested pattern (i.e., the impact of cognition increases, whereas the impact of affect decreases) is more pronounced in the case of consistent experiences than in the case of inconsistent experiences

A Longitudinal Experimental Study

Overview

To enhance the realism of the study, we placed the "customers" in an actual consumption situation. Specifically, we asked them to evaluate a newly created CD-ROM tutorial, which could be used and purchased to provide academic assistance in a difficult pricing class. We gave them three sample chapters (trials) of the CD-ROM tutorial over time in a computer-based format. After each trial, we asked them to solve a pricing problem related to the sample chapter, and then we gave them feedback on their performance. We manipulated performance by varying the quality of the tutorial and the feedback. After we provided the feedback, we elicited responses to key measures.

Sample

Our sample consisted of 157 marketing students enrolled in a graduate-level pricing class at a large German university. Because students represent the target market for the CD-ROM study guide, this sample is particularly relevant and appropriate. Participants were aware that previous students had experienced difficulties in this pricing class.

Research Design

The research design consisted of an 8 (levels of performance) × 3 (trial) full factorial design. Performance was a between-subjects factor, and trial was a within-subjects factor. To manipulate performance, we gave participants a sample chapter and asked them to solve a related pricing problem. In the high-performance condition, the content of the CD-ROM sample chapter made it easy to understand and to solve the pricing problem. This was accomplished by providing the different steps necessary to complete the pricing task successfully in a logical order. Furthermore, participants in this condition received positive feedback that the CD-ROM helped them solve the problem. Note that this feedback was designed to ensure that participants focused on the performance of the CD-ROM rather than on their own performance.

In the low-performance condition, the content of the CD-ROM chapter was not in logical order and therefore was difficult to read. It also contained little information related to the pricing problem. We also provided participants with negative feedback that the CD-ROM did not help them solve the pricing problem. Again, the goal was to focus participants on the performance with the CD-ROM rather than on their own performance.

To examine the proposed dynamic relationships, we manipulated performance across three different trials. In each trial, we presented participants with a different chapter from the CD-ROM tutorial and then asked them to solve a related pricing problem. Note that the three sample chapters covered completely different pricing issues and did not build on one another. Thus, learning effects, in the sense that knowledge gained in previous chapters could be helpful for understanding subsequent chapters and pricing tasks, were not present. In the high-performance condition, participants received the high-performance manipulation on all three trials. In the low-performance condition, participants received the low-performance manipulation on all three trials. These two conditions represent consistent performance experiences. We then manipulated intermediate levels of performance through high- and low-performance combinations across the three trials. To represent inconsistent experiences, we constructed two conditions with a performance experience on each subsequent trial that was different from the most recent one (+ – + and – + –). We outline the different conditions in Table 1.

Experimental Procedure

Before the first trial, we gave participants the introductory section, which described the purpose of the CD-ROM and content. We told them that a CD-ROM study guide had been developed to assist participants in solving difficult pricing problems in the course. Furthermore, we informed them that they would have the chance to test the CD-ROM tutorial before deciding if they wanted to purchase it. We then set up expectations about the CD-ROM tutorial and held them constant across the experimental conditions. We told participants that the CD-ROM tutorial contained 73 chapters, which would be similar to the ones they received in the testing phase but would cover different pricing topics. The introductory section of the CD-ROM tutorial informed participants that the purpose of the study guide was to help course participants understand difficult material in the class, and it provided an overview of the content.

Participants then received one sample chapter of the CD-ROM study guide (first trial), after which they solved a problem related to the material. We manipulated performance in the manner described previously. The measurement of cognition and affect occurred immediately after participants received the feedback from the performance manipulation. We then asked intervening questions to distract participants from the evaluation situation, after which we measured customer satisfaction. After the first trial, participants completed an intervening task. This involved reading a newspaper article about a recent pricing problem in practice and answering some open-ended questions about the content of the article.

TABLE 1
Overview of Experimental Conditions

Experimental Condition	Trial 1	Trial 2	Trial 3
1	+	+	+
2	+	+	–
3	+	–	+
4	+	–	–
5	–	+	+
6	–	+	–
7	–	–	+
8	–	–	–

Notes: + = high performance, and – = low performance.

We conducted the second and third trials in a similar manner. We told participants that they had the chance to get more experience with the CD-ROM study guide. They then received another sample chapter of the CD-ROM, solved a pricing task, and received feedback. After we measured the key variables, we distracted participants between the second and the third trial with an intervening task similar to the one described previously. After the third trial, participants completed a final set of general questions and provided demographic and classification information.

Measurement of Variables

The key variables of interest were cognition and affect. We measured cognition in terms of disconfirmation, which is the key cognitive component in the confirmation/disconfirmation paradigm (Oliver 1980). We measured disconfirmation with the following item, which parallels previous approaches in measuring disconfirmation (Fornell et al. 1996): "The performance of the CD-ROM meets my expectations." We evaluated the item on a seven-point Likert-type scale anchored by "I totally disagree" (1) and "I totally agree" (7).

We measured affect with three items. These items included elation, delight, and joy, which have also been measured in other studies examining affect (e.g., Holbrook and Batra 1987; Oliver 1993; Westbrook 1987; Westbrook and Oliver 1991). We asked participants how intensively they experienced these types of affect on a seven-point Likert-type scale.

In addition, we assessed affect and cognition with an open-ended, verbal protocol measures. We asked participants to describe any feelings and/or thoughts they had while using the tutorial. Responses were then coded into affective and cognitive categories. Responses were coded as affective if the statements related to any type of affect or feelings; they were coded as cognitive if they related to the content and functional performance of the tutorial. Research assistants carried out the codings. In terms of coder reliability, the coders agreed in 96% of the cases, and they resolved disagreements through discussions.

The key dependent variable was customer satisfaction, which we measured with the following three items: "All in all, I am satisfied with the CD-ROM tutorial"; "The CD-ROM tutorial compares to an ideal CD-ROM tutorial"; and "Overall, how satisfied are you with the CD-ROM tutorial?" We measured the items on an 11-point Likert-type scale. For the first two items, the scale ranged from "strongly agree" to "strongly disagree," and for the third item, it ranged from "very satisfied" to "very dissatisfied." The internal consistency of the satisfaction scale was excellent across the three trials (first trial: Cronbach's $\alpha = .95$; second trial: Cronbach's $\alpha = .97$; and third trial: Cronbach's

$\alpha = .96$). Thus, for further analyses, we calculated the satisfaction scores as the means of the satisfaction scale items.

Finally, we collected several variables as possible covariates. These included participants' age, gender, income, degree to which the budget for studying material is exhausted, perceived pressure to buy the CD-ROM tutorial, price consciousness, value consciousness, and self-confidence. Analyses indicate that the fixed effects of these covariates are not significant.[1] Thus, we dropped them from further analysis.

Results

H_1 predicted that the variance in customer satisfaction jointly explained by cognitive and affective factors increases as experience accumulates. As participants gain more experience with the CD-ROM tutorial across the three trials, the satisfaction judgment becomes more cumulative. To test H_1, we estimated the following regression model for each of the three trials:

$$(1) \qquad CS = b_0 + b_1 \times CO + b_2 \times AF + r,$$

where CS is customer satisfaction, CO is the cognitive predictor, AF is the affective predictor, and r is the random error.

We estimated the three regression models with the maximum likelihood method using the procedure MIXED in SAS 8.02.[2] The results appear in Table 2. First, in each of the three trials, there is a positive and statistically significant relationship between the cognitive predictor and satisfaction, and there is a positive and statistically significant relation between the affective predictor and satisfaction. Second, the results indicate that the R-square values substantially increase from the first (.645) to the third (.731) trial. These results support H_1.

Our second analysis pertained to the changing role of cognitive and affective factors over time. H_{2a} predicted that the impact of affective factors decreases as experience accumulates, and H_{2b} predicted that the role of cognitive factors increases as experience accumulates. To conduct a test of these hypotheses, we estimated the following regression model with interaction effects:

$$(2) \qquad CS_t = b_0 + b_1 \times CO_t + b_2 \times AF_t + b_3 \times time_t + b_4 \times CO_t \times time_t + b_5 \times AF_t \times time_t + r_t,$$

where CS_t is customer satisfaction, CO_t is cognition, and AF_i is affect at $time_t$. The different points in time are represented by the different trials (e.g., trial 1 = time 1). The random error associated with $time_t$ is represented by r_t.

In the testing of H_2, the coefficients of the two interaction terms, b_4 and b_5, are of key interest. These interaction terms are similar to slope-shift parameters (Mittal, Katrichis, and Kumar 2001), which enable us to estimate the differential slopes over time as experience accumulates. A significant, positive coefficient for the interaction term implies that the slope is higher for those who have more experience with the CD-ROM. In other words, the corresponding factor has a larger weight for customers who have more experience with the CD-ROM than for those who have less experience with it. On the basis of our theoretical reasoning, we expect the coefficient b_4 to be positive.

A negative coefficient for the interaction term implies that the slope is lower for those who have more experience with the CD-ROM. The corresponding factor has a smaller weight for customers who have more experience with the CD-ROM than for those who have less experience with it. On the basis of our theoretical reasoning, we expect the coefficient b_5 to be negative. Furthermore, statistically nonsignificant coefficients imply that the weight of the corresponding factor is invariant to the experience a customer has with the CD-ROM. We estimated the model with the maximum likelihood method using the procedure MIXED in SAS 8.02. The estimation results appear in Table 3.

The results indicate that there is an overall statistically significant, positive relationship between cognition and satisfaction ($b_1 = 22.380$, $p < .001$) and between affect and satisfaction ($b_2 = 8.490$, $p < .001$). In addition, there is a significant, positive interaction between cognition and time ($b_4 = 1.960$, $p < .05$). This means that the impact of cognitive factors on the satisfaction judgment increases over time. Thus, there is support for H_{2a}.

Furthermore, there is a significant, negative interaction between affect and time ($b_5 = -2.330$, $p < .05$), supporting the notion that the impact of affective factors on customer satisfaction decreases over time. Thus, there is also support for H_{2b}. Taking these two findings together, the impact of cognitive factors on the satisfaction judgment increases over time, whereas the impact of affective factors decreases.

Further support for these hypotheses is provided by examining the results for the verbal protocol measures. As we mentioned previously, open-ended responses were coded into affective and cognitive categories. Overall, there were more cognitive statements than affective ones. We expected this because the product category is cognitive in nature. More important, however, the pattern of responses over the trials was consistent with the hypotheses. Specifically, the largest number of emotional statements was made in the first trial (23) and then decreased (14 in both the second and the third trials). For cognitive statements related to the content and functional performance, participants exhibited an increasing trend. On the first trial, 74 such statements were made, followed by 89 in the second trial and 90 in the third.

H_3 proposed that an inconsistent performance experience would attenuate the increasing impact of cognitive factors and the decreasing impact of affective factors over time

---

[1] The individual significance levels were as follows: age: $p = .485$; gender: $p = .140$; income: $p = .825$; degree to which the budget for studying material is exhausted: $p = .354$; perceived pressure to buy the CD-ROM tutorial: $p = .563$; price consciousness: $p = .223$; value consciousness: $p = .295$; and self-confidence: $p = .103$.

[2] We used full maximum likelihood estimation because the focus of the analysis is on deviance tests and not on the random part parameters, for which the restricted maximum likelihood method is preferable (Singer and Willett 2003; Snijders and Bosker 1999).

TABLE 2
Regression Results Across Trials

		Customer Satisfaction Trial 1			Customer Satisfaction Trial 2			Customer Satisfaction Trial 3		
−2 log-likelihood		940.9			525.8			499.7		
R^2		.645			.706			.731		
		Solutions for Fixed Effects								
Parameter	**Effect**	**Estimate**	**t-Value**	***p***	**Estimate**	**t-Value**	***p***	**Estimate**	**t-Value**	***p***
b_0	Intercept	6.318	58.940	.000	6.074	58.950	.000	5.823	61.400	.000
b_1	Cognition	.682	10.510	.000	.787	12.890	.000	.849	15.870	.000
b_2	Affect	.353	6.060	.000	.303	5.160	.000	.171	3.340	.001
		Solutions for Residual Effects								
Parameter		**Estimate**	**Z Value**	***p***	**Estimate**	**Z Value**	***p***	**Estimate**	**Z Value**	***p***
σ^2 (variance of r_{ij})		1.804	8.800	.000	1.667	8.860	.000	1.412	8.860	.000

(compared with a consistent performance experience). To test this hypothesis, we created two subgroups from the data set. The first group had consistent satisfaction/performance experiences (+ + + or – – –), whereas the second group had completely inconsistent performance experiences (+ – + or – + –).

We then ran the regression model with interaction effects described previously for each group separately. As we show in Table 4, for the consistent performance experiences, the impact of cognition increases slightly, and the impact of affect decreases over time, as is evidenced by the results for the interaction effects ($b_4 = .119$, $p < .06$; $b_5 = –.156$, $p < .05$). In the case of inconsistent performance experiences, however, neither of these interaction terms is significant (see Table 4). In other words, in the case of inconsistent performance experiences, there is neither a significant increase of cognition over time nor a significant decrease of affect over time in predicting customer satisfaction. Thus, inconsistent performance experiences attenuate these effects. These results support H_3.

Discussion

The goal of this study was to examine the joint effects of cognitive and affective factors on satisfaction judgments in a dynamic setting. The results of our study show that together, cognitive and affective factors explain the variance in satisfaction judgments well and that the strength of association increases over time. More important, we found that as the number of experiences increases over time, the influence of cognitive factors increases, whereas the influence of affective factors decreases. However, this effect attenuates with inconsistent consumption experiences. These findings have several theoretical and managerial implications.

TABLE 3
Regression Results with Interaction Effects

–2 log-likelihood	1566.9			
		Solution for Fixed Effects		
Parameter	**Effect**	**Estimate**	**t-Value**	***p***
b_0	Intercept	6.072	103.200	.00
b_1	Cognition	.773	22.380	.00
b_2	Affect	.275	8.490	.00
b_3	Time	–.248	–3.440	.00
b_4	Cognition × time	.083	1.960	.05
b_5	Affect × time	–.091	–2.330	.02
		Solution for Residual Effect		
Parameter		**Estimate**	**Z Value**	***p***
σ^2		1.630	15.350	.00

TABLE 4
Regression Results for Consistent and Inconsistent Performance Experience

		Consistent Performance Experience			Inconsistent Performance Experience		
–2 log-likelihood		449.5			422.8		
		Solution for Fixed Effects			Solution for Fixed Effects		
Parameter	**Effect**	**Estimate**	**t-Value**	***p***	**Estimate**	**t-Value**	***p***
b_0	Intercept	6.095	59.660	.00	6.268	48.320	.00
b_1	Cognition	.742	14.610	.00	.833	10.450	.00
b_2	Affect	.335	6.600	.00	.249	8.490	.00
b_3	Time	–.197	–1.590	.12	–.119	3.220	.45
b_4	Cognition × time	.119	1.870	.06	.048	–.750	.62
b_5	Affect × time	–.156	–2.500	.01	.111	1.200	.23
		Solution for Residual Effect			Solution for Residual Effect		
Parameter		**Estimate**	**Z Value**	***p***	**Estimate**	**Z Value**	***p***
σ^2		1.419	15.350	.00	1.9844	15.350	.00

Implications

Several implications follow from our study. Most important, our study contributes to the existing literature on customer satisfaction. First, we illustrate the importance of examining customer satisfaction from a dynamic perspective. We show that the antecedents leading to customer satisfaction vary over time. As we mentioned previously, it is well recognized that both cognition and affect have important influences on the satisfaction judgment. However, the key finding of this study points to the differential roles of these factors as experience accumulates. To the best of our knowledge, this is the first study to examine the role of cognition and affect simultaneously in a dynamic setting. Thus, we provide a deeper integrative understanding of how cognition and affect combine to form the customer satisfaction judgment. Our findings show that these judgments are based on different factors, depending on the amount of experience that has been accumulated over time. Thus, capturing satisfaction judgments at only one time may not provide a full picture of the underlying process.

Second, most studies in this area examine in a cross-sectional manner the satisfaction judgment process after customers already have some experience or knowledge with the product or service. In the current study, we investigated the customer satisfaction formation process beginning with the very first consumption experience. This provides a deeper insight into how the satisfaction judgment crystallizes over time. In this context, we found that affective factors have their strongest role in the early stages. The cognitive component maintains its importance and increases its role over time.

Third, we examine the role of cognition and affect in influencing satisfaction judgment processes. Our findings suggest that affect is particularly important in the early stages of the judgment formation process in which customers have little knowledge or experience related to the product. In other words, when a stimulus is new and unfamiliar, feelings appear to play a critical role in the construction of satisfaction judgments, consistent with the how-do-I-feel-about-it heuristic (Gorn, Goldberg, and Basu 1993; Pham 1998; Schwartz and Clore 1988). As experience accumulates, however, the role of cognition becomes more prominent. That is, a direct access strategy (Forgas 1994, 1995) can be employed in which evaluations from previous trials (based on a more detailed evaluation of product attributes) can be recalled from memory.

Fourth, we provide insights into how the inconsistency versus consistency of performance feedback can also affect the customer satisfaction formation process. Specifically, when experiences are inconsistent, the related judgments are less stable and predictable. In this case, the changing patterns of affect and cognition do not emerge. An implication of our findings is that affective factors are particularly important and relevant when the judgments are based on little information (i.e., they are newly formed) or when information is inconsistent. In these cases, the cognitive information has unstable predictive validity, so customers are less likely to rely on this information. This finding emphasizes the importance of identifying how the pattern or nature of experiences can influence satisfaction judgments as well as the importance of managing customer experiences. If customers have inconsistent experiences, their evaluations of the product/service will be less crystallized and less stable. Thus, in an effort to improve customer satisfaction and loyalty, it is critical to ensure that customer experiences are positive and consistent.

Fifth, the findings provide further support for research on attitude strength in an alternative context (Jaccard et al. 1995; Jacoby et al. 2002; Petty and Krosnick 1995). That is, cognitive and affective factors were better able to predict satisfaction evaluations as the number of trials increased, particularly in the case of consistent experiences. This is in line with the notion that newly formed attitudes are less stable and, therefore, less predictive of information processing and behavior (Eagly and Chaiken 1993). As knowledge and experience accumulates, the relationship between the cognitive and affective factors and satisfaction is strengthened.

Sixth, our study also makes a methodological contribution by developing a procedure for studying these dynamic effects in a tightly controlled, yet realistic setting. This methodology could be employed to examine judgment formation processes in various other product/service contexts as well.

Finally, on a general level, our study suggests that it is insightful to investigate marketing phenomena from a dynamic perspective. As our study shows, judgment patterns change over time.

Limitations and Avenues for Further Research

Several limitations of this study are worth mentioning. In turn, these provide avenues for further research. First, the study was based on a highly utilitarian product (i.e., a CD-ROM tutorial). In general, the consumption of utilitarian goods is highly cognitively driven and accomplishes a functional or practical task (Strahilevitz and Myers 1998). This may partially explain why the role of cognition (affect) became stronger (weaker) over trials. In other words, as customers gained experience in evaluating the utilitarian benefits of the product, these cognitive factors gained importance in influencing satisfaction. However, it is not clear whether the same pattern would occur for highly hedonic products/services. The consumption of hedonic products/services is characterized by highly affective and sensory experiences of sensual and/or aesthetic pleasure (Dhar and Wertenbroch 2000; Hirschman and Holbrook 1982; Okada 2005). Thus, in such cases, the impact of affect might not decline over time. It is possible that a direct access strategy will still be employed as experience accumulates (i.e., a recall of a previous crystallized evaluation from memory), but this prior evaluation would be based more on affective than on cognitive factors. In support of this view, Kempf (1999) finds that for hedonic products in particular, affective responses to the trial are powerful antecedents of consumers' evaluations of a trial experience and, subsequently, brand attitude. In addition, it might be the case that the relative importance of cognitive versus affective drivers of customer satisfaction varies across product categories. It is plausible to assume that the relative importance of cognitive factors is higher for utilitarian products than for hedonic ones. These speculations, which are beyond the scope of this article, show that a more detailed investigation of affective and cognitive drivers of customer satisfaction that accounts for different product categories is a fruitful avenue for further research.

Second, in addition to the product category type, customers' product involvement could also play an important role in this context. Product involvement can be defined as the perceived personal relevance of a product and a customer's inherent interest in a product (Wilkie 1994; Zaichkowsky 1985). This study investigated the proposed relationships in a high-involvement context. In these situations, customers usually employ a more analytical information-processing strategy and exhibit increased cognitive elaborations (Petty, Cacioppo, and Schumann 1983; Wright 1973). Thus, they are likely to base their decisions more on cognitive than on affective factors. Conversely, in low-involvement situations, information processing is usually more holistic, and thus decisions might be influenced more by affective factors. Apart from these effects, it might be interesting to investigate whether and how these potential effects of customers' product involvement interact with the product category type (hedonic and utilitarian). Thus, studying the relevance of involvement in the context of affective and cognitive drivers of customer satisfaction is another interesting area for further research.

Third, in our empirical study, we measured only the prevalence of positive affect and did not consider negative affect. Because it has been shown that positive and negative affect can have asymmetric effects on customers' judgments and choices (Mittal, Ross, and Baldasare 1998), a promising avenue for further research might be to investigate positive and negative affect simultaneously. Prior research also demonstrates that different types of (negative) emotions can have opposing moderating effects (Garg, Inman, and Mittal 2005). Thus, additional research on the role of cognition and affect in the formation of customer satisfaction might further differentiate between different types of emotions.

Implications for Marketing Practice

Our findings also have several important implications for marketing managers. Our study helps managers understand customer satisfaction in a more thorough way. It sheds light on the formation process of customer satisfaction, and managers learn that customer satisfaction has a more stochastic character in the early stages. This gives managers the opportunity to influence the satisfaction judgment to a greater extent in the early stages because the satisfaction has not been crystallized.

Furthermore, it is common in practice for managers to consider customer satisfaction in a logical, rational manner (i.e., if the product or service performs well, satisfaction will be higher). The results of our study point out that affective factors can play a critical role as well, particularly in the early stages of the satisfaction formation process. Thus, achieving customer satisfaction involves not merely disconfirmation judgments but a subjective, affective component as well. This point is particularly important when companies are in the early stages of establishing a relationship with a customer. Thus, for new relationships or new products, managers must pay close attention to affective aspects and be careful to manage them effectively.

Moreover, our study provides insights into managing customer satisfaction. For example, it shows how important it is to manage affect when customers have inconsistent experiences over time. Thus, if companies are trying to recover from a situation in which their products/services have suffered from inconsistent performance, it appears that it is important to manage the affective factors surrounding the satisfaction in addition to product/service quality considerations. In addition, providing customers with consistent, positive experiences is critical to the development of stable satisfaction judgments.

REFERENCES

Alvero, Alicia M., Barbara R. Bucklin, and John Austin (2001), "An Objective Review of the Effectiveness and Essential Characteristics of Performance Feedback in Organizational Settings," *Journal of Organizational Behavior Management*, 21 (1), 3–29.

Anderson, Eugene W., Claes Fornell, and Donald R. Lehmann (1994), "Customer Satisfaction, Market Share, and Profitability: Findings from Sweden," *Journal of Marketing*, 58 (July), 53–66.

———, ———, and Roland T. Rust (1997), "Customer Satisfaction, Productivity, and Profitability: Differences Between Goods and Services," *Marketing Science*, 16 (2), 129–45.

——— and Mary W. Sullivan (1993), "The Antecedents and Consequences of Customer Satisfaction for Firms," *Marketing Science*, 12 (2), 125–43.

Bearden, William O. and Jesse E. Teel (1983), "Selected Determinants of Consumer Satisfaction and Complaint Reports," *Journal of Marketing Research*, 20 (February), 21–28.

Berscheid, Ellen, William Graziano, Thomas Monson, and Marshall Dermer (1976), "Outcome Dependency: Attention, Attribution, and Attraction," *Journal of Personality and Social Psychology*, 34 (3), 978–89.

Bettman, James R., Mary Frances Luce, and John W. Payne (1998), "Constructive Consumer Choice Processes," *Journal of Consumer Research*, 25 (3), 187–217.

Bolton, Ruth N. (1998), "A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction," *Marketing Science*, 17 (1), 45–65.

——— and James H. Drew (1991), "A Longitudinal Analysis of the Impact of Service Changes on Customer Attitudes," *Journal of Marketing*, 55 (January), 1–9.

——— and Katherine N. Lemon (1999), "A Dynamic Model of Customers' Usage of Services: Usage as an Antecedent and Consequence of Satisfaction," *Journal of Marketing Research*, 36 (May), 171–86.

Boulding, William, Ajay Kalra, Richard Staelin, and Valarie A. Zeithaml (1993), "A Dynamic Process Model of Service Quality: From Expectations to Behavioral Intentions," *Journal of Marketing Research*, 30 (February), 7–27.

Chandrashekaran, Murali, Kevin McNeilly, Frederick A. Russ, and Detelina Marinova (2000), "From Uncertain Intentions to Actual Behavior: A Threshold Model of Whether and When Salespeople Quit," *Journal of Marketing Research*, 37 (November), 463–79.

Dhar, Ravi and Klaus Wertenbroch (2000), "Consumer Choice Between Hedonic and Utilitarian Goods," *Journal of Marketing Research*, 37 (February), 60–71.

Dover, Philip A. and Jerry C. Olson (1977), "Dynamic Changes in an Expectancy-Value Attitude Model as a Function of Multiple Exposures to Product Information," in *Contemporary Marketing Thought: Proceedings of the 1977 AMA Conference*, Barnett A. Greenberg and Danny N. Bellenger, eds. Chicago: American Marketing Association, 455–60.

Eagly, Alice H. and Shelly Chaiken (1993), *The Psychology of Attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Farley, John U., Jerrold P. Katz, and Donald R. Lehmann (1978), "Impact of Different Comparison Sets on Evaluation of a New Subcompact Car Brand," *Journal of Consumer Research*, 5 (September), 138–42.

Fazio, Russel H. (1986), "How Do Attitudes Guide Behavior?" in *Handbook of Motivation and Cognition: Foundations of Social Behavior*, Richard M. Sorrentino and Edward T. Higgins, eds. Chichester. John Wiley & Sons, 204–243.

Fiedler, Klaus (1991), "On the Task, the Measures and the Mood: Research on Affect and Social Cognition," in *Emotion and Social Judgments*, Joseph P. Forgas, ed. Elmsford, NY: Pergamon Press, 83–104

Forgas, Joseph P. (1994), "The Role of Emotion in Social Judgments: An Introductory Review and an Affect Infusion Model (AIM)," *European Journal of Social Psychology*, 24 (1), 1–24.

——— (1995), "Mood and Judgment: The Affect Infusion Model (AIM)," *Psychological Bulletin*, 117 (1), 39–66.

Fornell, Claes (1992), "A National Customer Satisfaction Barometer: The Swedish Experience," *Journal of Marketing*, 56 (January), 6–21.

———, Michael D. Johnson, Eugene W. Anderson, Jaesung Cha, and Barbara Everitt Bryant (1996), "The American Customer Satisfaction Index: Nature, Purpose, and Findings," *Journal of Marketing*, 60 (October), 7–18.

Fournier, Susan and David G. Mick (1999), "Rediscovering Satisfaction," *Journal of Marketing*, 63 (October), 5–23.

Garg, Nitika, J Jeffrey Inman, and Vikas Mittal (2005), "Incidental and Task-Related Affect: A Re-Inquiry and Extension of the Influence of Affect on Choice," *Journal of Consumer Research*, 32 (June), 154–59.

Gorn, Gerald J., Marvin E. Goldberg, and Kunal Basu (1993), "Mood, Awareness, and Product Evaluation," *Journal of Consumer Psychology*, 2 (3), 237–56.

Heslin, Richard, Brian Blake, and James Rotton (1972), "Information Search as a Function of Stimulus Uncertainty and the Importance of the Response," *Journal of Personality and Social Psychology*, 23 (3), 333–39.

Hirschman, Elizabeth C and Morris B. Holbrook (1982), "Hedonic Consumption· Emerging Concepts, Methods, and Propositions," *Journal of Marketing*, 46 (Summer), 92–101.

Holbrook, Morris B. and Rajeev Batra (1987), "Assessing the Role of Emotions as Mediators of Consumer Responses to Advertising," *Journal of Consumer Research*, 14 (December), 404–420.

Jaccard, James, Carmen M. Radecki, Tracey E. Wilson, and Patricia J. Dittus (1995), "Methods for Identifying Consequential Beliefs: Implications for Understanding Attitude Strength," in *Attitude Strength: Antecedents and Consequences*, Richard E. Petty and Jon A. Krosnick, eds. Mahwah, NJ: Lawrence Erlbaum Associates, 337–59.

Jacoby, Jacob, Maureen Morrin, James Jaccard, Zeynep Gurhan, Alfred Kuss, and Durairaj Maheswaran (2002), "Mapping Attitude Formation as a Function of Information Input: Online Processing Models of Attitude Formation," *Journal of Consumer Psychology*, 12 (1), 21–34.

Kahneman, Daniel and A. Tversky (1973), "On the Psychology of Prediction," *Psychological Review*, 80 (4), 237–51.

Kempf, DeAnna S. (1999), "Attitude Formation from Product Trial: Distinct Roles of Cognition and Affect for Hedonic and Functional Products," *Psychology & Marketing*, 16 (1), 35–50.

LaBarbera, Priscilla A. and David Mazursky (1983), "A Longitudinal Assessment of Consumer Satisfaction/Dissatisfaction: The Dynamic Aspect of the Cognitive Process," *Journal of Marketing Research*, 20 (November), 393–404.

Mano, Haim and Richard L. Oliver (1993), "Assessing the Dimensionality and Structure of the Consumption Experience: Evaluation, Feeling, and Satisfaction," *Journal of Consumer Research*, 20 (3), 451–67.

Mittal, Vikas, Jerome M. Katrichis, and Pankaj Kumar (2001), "Attribute Performance and Customer Satisfaction over Time: Evidence from Two Field Studies," *Journal of Services Marketing*, 15 (5), 343–56.

———, Pankaj Kumar, and Michael Tsiros (1999), "Attribute-Level Performance, Satisfaction, and Behavioral Intentions over Time: A Consumption-System Approach," *Journal of Marketing*, 63 (April), 88–101.

———, William T. Ross, and Patrick M. Baldasare (1998), "The Asymmetric Impact of Negative and Positive Attribute-Level Performance on Overall Satisfaction and Repurchase Intentions," *Journal of Marketing*, 62 (January), 33–47.

Okada, Erica M. (2005), "Justification Effects on Consumer Choice of Hedonic and Utilitarian Goods," *Journal of Marketing Research*, 42 (February), 43–53.

Oliver, Richard L. (1980), "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research*, 17 (November), 460–69.

——— (1993), "Cognitive, Affective, and Attribute Bases of Satisfaction Response," *Journal of Consumer Research*, 20 (December), 418–30.

——— (1997), *Satisfaction: A Behavioral Perspective on the Consumer*. Boston· McGraw-Hill.

——— and Wayne S. DeSarbo (1988), "Response Determinants in Satisfaction Judgments," *Journal of Consumer Research*, 14 (March), 495–507.

——— and John E. Swan (1989a), "Consumer Perceptions of Interpersonal Equity and Satisfaction in Transactions· A Field Survey Approach," *Journal of Marketing*, 53 (April), 21–35.

——— and ——— (1989b), "Equity and Disconfirmation Perceptions as Influences on Merchant and Product Satisfaction," *Journal of Consumer Research*, 16 (December), 372–83.

Petty, Richard E., John T. Cacioppo, and David T. Schumann (1983), "Central and Peripheral Routes to Advertising Effectiveness: The Moderating Effect of Involvement," *Journal of Consumer Research*, 10 (September), 135–46.

——— and Jon A. Krosnick (1995), *Attitude Strength. Antecedents and Consequences*. Mahwah, NJ: Lawrence Erlbaum Associates.

———, H. Rao Unnava, and Alan J. Strathman (1991), "Theories of Attitude Change," in *Handbook of Consumer Behavior*, Thomas S. Robertson and Harold H. Kassarjian, eds. Englewood Cliffs, NJ: Prentice Hall, 241–80.

Pham, Michel Tuan (1998), "Representativeness, Relevance, and the Use of Feelings in Decision Making," *Journal of Consumer Research*, 25 (September), 144–60.

———, Joel B. Cohen, John W. Pracejus, and G. David Hughes (2001), "Affect Monitoring and the Primacy of Feelings in Judgment," *Journal of Consumer Research*, 28 (September), 167–88.

Rust, Roland T., Christine Moorman, and Peter R. Dickson (2002), "Getting Return on Quality: Revenue Expansion, Cost Reduction, or Both?" *Journal of Marketing*, 65 (October), 7–24.

——— and Anthony J. Zahorik (1993), "Customer Satisfaction, Customer Retention, and Market Share," *Journal of Retailing*, 69 (2), 193–215.

Schwartz, Nobert and Gerald L. Clore (1988), "How Do I Feel About It? Informative Functions of Affective States," in *Affect, Cognition, and Social Behavior*, Klaus Fielder and Joseph Forgas, eds. Toronto: Hogrefe International, 44–62.

Singer, Judith D. and John B. Willett (2003), *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. New York: Oxford University Press.

Slotegraaf, Rebecca J. and J. Jeffrey Inman (2004), "Longitudinal Shifts in the Drivers of Satisfaction with Product Quality: The Role of Attribute Resolvability," *Journal of Marketing Research*, 41 (August), 269–80.

Smith, Amy K. and Ruth N. Bolton (2002), "The Effect of Customers' Emotional Responses to Service Failures on Their Recovery Effort Evaluations and Satisfaction Judgments," *Journal of the Academy of Marketing Science*, 30 (1), 5–23.

Smith, Robert E. and William R. Swinyard (1983), "Attitude–Behavior Consistency: The Impact of Product Trial Versus Advertising," *Journal of Marketing Research*, 20 (August), 257–67.

Snijders, Tom A.B. and Roel J. Bosker (1999), *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. London. Sage Publications.

Stone, Dianna L. and Eugene F. Stone (1985), "The Effects of Feedback Consistency and Feedback Favorability on Self-Perceived Task Competence and Perceived Feedback Accuracy," *Organizational Behavior and Human Decision Processes*, 36 (2), 167–85.

Strahilevitz, Michal and John Myers (1998), "Donations to Charity as Purchase Incentives: How Well They Work May Depend on What You Are Trying to Sell," *Journal of Consumer Research*, 24 (2), 434–46.

Szymanski, David M. and David Henard (2001), "Customer Satisfaction: A Meta-Analysis of the Empirical Evidence," *Journal of the Academy of Marketing Science*, 29 (1), 16–35.

Westbrook, Robert A. (1987), "Product/Consumption-Based Affective Responses and Postpurchase Processes," *Journal of Marketing Research*, 24 (August), 258–70.

——— and Richard L. Oliver (1991), "The Dimensionality of Consumption Emotion Patterns and Consumer Satisfaction," *Journal of Consumer Research*, 18 (June), 84–91.

Wilkie, William L. (1994), *Consumer Behavior*, 3d ed. New York: John Wiley & Sons.

Wright, Peter L. (1973), "The Cognitive Processes Mediating Acceptance of Advertising," *Journal of Marketing Research*, 10 (February), 53–62.

Zaichkowsky, Judith L. (1985), "Measuring the Involvement Construct," *Journal of Consumer Research*, 12 (December), 341–52.

Anick Bosmans

Scents and Sensibility: When Do (In)Congruent Ambient Scents Influence Product Evaluations?

This research examines when and how ambient scents (i.e., scents that are present in the environment and do not emanate from a specific product) affect product evaluations. Consistent with theory in environmental psychology, the author predicts that such ambient scents can elicit emotional responses that can influence subsequent judgments, but only when specific conditions are met. Three factors are hypothesized to moderate the effects of pleasant ambient scents on evaluations: (1) the congruence of the scent with the product, (2) the salience of the scent, and (3) consumers' motivation to correct for extraneous influences. The findings of three experiments suggest that as long as ambient scents are congruent with the product, they continue to affect consumers' evaluations, even when their influence becomes salient or when consumers are sufficiently motivated to correct for extraneous influences. In addition, as long as the scent is not completely incongruent with the product, noncongruent scents are corrected for only when their influence becomes salient or when consumers are sufficiently motivated. These findings indicate that the impact of pleasant ambient scents on product evaluations may be stronger than that of other affective environmental cues.

Retailers have been using scents in the marketplace for a long time to influence consumers' buying behavior and satisfaction. The air from an in-store bakery is often blown back into the store to increase consumers' interest in that department and to signal freshness and quality. In the past, retailers were limited to using product-specific scents emanating from the products themselves, but recent technologies now enable retailers to apply a diverse selection of (artificial) ambient scents into the retail environment. For example, British Airways distributes a fragrance called Meadow Grass into the air of its business lounges to make customers feel comfortable. Similarly, Thomas Pink, an upscale British shirt retailer, uses sensors that emit the smell of freshly laundered cotton (Ellison and White 2000).

Contrary to product-specific scents, ambient scents do not emanate from a specific product and therefore can be considered extraneous environmental cues. Ambient scents are not necessarily related to specific products, and their possible influence on consumer behavior goes beyond that of signalling mere product quality. Whereas the effects of product-specific scents are well understood (see, e.g., Bone and Jantrania 1992), much less is known about how ambient scents influence product evaluations and whether consumers can control for their extraneous influence. For example, consider a supermarket that distributes the famous fragrance Channel No. 5 to create a pleasant store atmosphere. Imagine this fragrance reaching the fresh vegetables and fruit department. How would this ambient scent affect consumers' product evaluations?

Despite marketing practitioners' increased interest in the use of ambient scents and despite the technological progress in scent applications, research investigating the impact of ambient scents on product evaluations is limited. In addition, the few studies that have examined the impact of ambient scents on consumers' judgments have yielded mixed results (Bone and Ellen 1999): Although pleasant scents sometimes increase evaluations (Spangenberg, Crowley, and Henderson 1996), they do not always seem to do so (Morrin and Ratneshwar 2003). It is unclear whether and to what extent ambient scents influence consumers' product evaluations.

This research examines the impact of ambient scents on product evaluations in more detail. Consistent with theory in environmental psychology, it is assumed that as is the case with other environmental stimuli, such as in-store music or colors, ambient scents elicit emotional responses that can influence product and store evaluations (Bitner 1992; Gorn, Pham, and Sin 2001; Zeithaml and Bitner 2000). In addition, because ambient scents are extraneous environmental cues, consumers are likely to correct for their possible erroneous influences. More specifically, it is proposed that correcting for the extraneous influence of ambient scents depends on at least three conditions: (1) how congruent the scent is with the product category, (2) how salient the extraneous influence is, and (3) consumers' motivation to correct for possible extraneous influences. The results of three experiments show that, indeed, consumers seem to correct for the extraneous influence of ambient

---

Anick Bosmans is an assistant professor, Marketing Department, Tilburg University, the Netherlands (e-mail: A.M.M.Bosmans@uvt.nl). The author is grateful to Rik Pieters and Dirk Smeesters for insightful comments on previous drafts of this article. Anouk Lepinoy's and Sandra Zenden's help in collecting the data is also gratefully appreciated. This research is financially supported by The Netherlands Organization for Health Research and Development (Grant No. 014-12-011).

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

scents only when they perceive the scent as incongruent with the product category. In addition, discounting the influence of incongruent ambient scents occurs only when the influence of the ambient scent is made salient to respondents (Experiment 1) or when consumers have sufficient processing motivation to correct for the extraneous influence (Experiments 2 and 3). Consumers are able to correct regardless of their processing motivation only when ambient scents are highly incongruent with the product category (Experiment 3).

Ambient Scents and Product Evaluations

Ambient scents can be considered affective cues that guide consumers' subsequent product evaluations. This assumption is consistent with the observation that people's primary reaction to scents is one of like or dislike. Indeed, the most salient attribute of a scent seems to be its pleasantness or unpleasantness (Engen 1982). Moreover, smell is directly linked to the experience of emotions (Herz and Engen 1996): The primary olfactory cortex forms a direct, anatomical link with the amygdala-hippocampal complex of the limbic system, and only two synapses separate the olfactory nerve from the amygdala, which is critical for the experience of an emotion (Aggleton and Mishkin 1986). No other sensory system makes this kind of direct and intense contact with the neural substrates of an emotion. As a result, olfaction is often considered *the* emotional sense (Ehrlichman and Halpern 1988; Engen 1982). Consistent with this notion, various streams of research have suggested that scents influence human behavior through the affect they elicit (Baron 1990; Baron and Bronfen, 1994; Baron and Thomley 1994; Knasko 1995; Rotton et al. 1978). For example, both task performance and helping behavior increase when a pleasant scent is present in the environment, and this effect seems to be mediated by people's affective reactions (Baron 1990; Baron and Thomley 1994).

Despite the ample evidence that pleasant ambient scents increase approach behavior, such as increased in-store lingering (Knasko 1989), increased time spent looking at advertisements (Knasko 1995), and increased amounts of money spent in slot machines (Hirsch 1995), little is known about how ambient scents influence product evaluations. Few studies have investigated the impact of ambient scents on evaluations of products, environments, or people. In addition, these studies have yielded mixed results (Bone and Ellen 1999). Although scents sometimes appear to affect evaluations (e.g., Rotton et al. 1978; Spangenberg, Crowley, and Henderson 1996), they do not always do so (Cann and Ross 1989; Morrin and Ratneshwar 2003). The purpose of the current research is to investigate the impact of pleasant ambient scents on product evaluations in further detail and, in particular, to test conditions in which scents affect these evaluations.

Consistent with the affect-as-information literature (Pham 1998; Schwarz and Clore 1983), it is argued that as is the case with other affective environmental stimuli, such as music or the weather, ambient scents can affect subsequent evaluations because people often misconsider their reactions toward the extraneous stimulus as a reaction toward the product under evaluation. Indeed, when consumers evaluate products, they often inspect their momentary feelings (i.e., "How do I feel about it?") and consider them informative for their evaluations (Pham 1998), even when these feelings originate from irrelevant sources, such as in-store music, the weather, or, in the current context, ambient scents. Thus, given the strong emotional potency of ambient scents, it is logical to assume that they can provide affective information (i.e., "I feel good about it, so the product must be good"), even when these scents do not arise from the product under evaluation.

Controlling for the Influence of Ambient Scents

In general, relying on extraneous affect when making judgments seems to be a spontaneous process, in that people appear to automatically consider their affect representative for the target (Martin 1986; Schwarz 1990). However, because most people like to render judgments that are free of extraneous influences, they may subsequently assess the appropriateness of their initial reaction (Wegener and Petty 2001).

This article examines how ambient scents affect subsequent product evaluations and when consumers control for these extraneous influences. On the basis of previous theorizations in the domain of contextual correction (Martin 1986; Schwarz and Bless 1992; Wegener and Petty 2001), it is argued that whether people control for the extraneous influence of an ambient scent depends on at least three conditions: (1) the scent's congruence with the product category, (2) the salience of the ambient scent, and (3) consumers' motivation to correct for possible extraneous influences.

Congruence of the Ambient Scent with the Product Category

It seems reasonable to assume that correction for the erroneous influences of ambient scents is more likely when these scents are perceived as inappropriate for the product evaluation. In addition to their level of pleasantness, ambient scents provide semantic information (Bone and Jantrania 1992; Mitchell, Kahn, and Knasko 1995). For example, the pleasant smell of lavender not only elicits pleasant feelings but also may activate memory concepts of a holiday in the French Provence. Similarly, the smell of roses may remind a person of baby care products and may activate baby-related memory concepts. As a result of these semantic connections, ambient scents can vary in their level of congruence with the product under evaluation.

Given these strong semantic connections, it can be expected that the effect of ambient scents on product evaluations depends on how congruent they are with the product. According to the contextual correction literature (Martin 1986; Schwarz and Bless 1992; Wegener and Petty 2001), the level of congruence between an extraneous cue and the target is indeed a determinant of whether correction will occur. When overlap is high, separating a person's reaction toward the context from his or her reaction toward the target is difficult because both the context and the target jointly activate similar memory concepts. As a result, correction is less likely when overlap is high. In addition, extraneous cues that are incongruent with the target are considered "inappropriate" and alert people to their potentially biasing influence (Campbell and Kirmani 2000; Wegener and Petty 2001). In contrast, cues that are congruent with the target are considered appropriate and therefore are less likely to promote correction processes.

Thus, it is hypothesized that when the semantic properties of an ambient scent are highly congruent with those of the product, the similarity between the product and the scent makes it less likely for consumers to correct for the erroneous influence of the extraneous scent. This hypothesis is consistent with recent findings from the affect-as-information literature that suggest that extraneous affective information that bears some domain resemblance with the target is more likely to influence judgments than affective information that is unrelated to the target (Bosmans and Baumgartner 2005; Raghunathan, Pham, and Corfman 2006). Thus, consumers are more likely to control for the influence of ambient scents when these scents are perceived as incongruent with the product.

Although this expectation makes sense from a theoretical perspective and is consistent with the existing literature on contextual correction (Martin 1986; Schwarz and Bless 1992; Wegener and Petty 2001), the ambient scent literature has been less supportive for the argument that congruent scents are more likely to affect product evaluations than incongruent scents (Knasko 1995; Morrin and Ratneshwar 2003). For example, people who were exposed to a chocolate scent evaluated a set of chocolate slides equally positive compared with people who were exposed to a baby powder scent (Knasko 1995). However, these studies appear to have investigated congruency effects in isolation and have not examined factors that may moderate the effects of congruent versus incongruent scents on product evaluations.

An incongruent scent does not necessarily mean that consumers automatically consider its influence extraneous to the judgmental process or that consumers correct for its erroneous influence. On the basis of several streams of literature (Martin, Seta, and Crelia 1990; Petty et al. 1993; Schwarz and Clore 1983), it is proposed that whether consumers control for the inappropriate influence of ambient scents depends on (1) how salient the ambient scent is and (2) how motivated consumers are to correct for erroneous influences. It is believed that both conditions make people more aware of the erroneous influence of the ambient scent.

Salience of Ambient Scents

In most cases, people consider their initial reaction a genuine affective response to the target (Pham 1998; Schwarz and Clore 1983). However, as the salience of the extraneous source increases, people become more aware that a source other than the target can be responsible for their initial reaction. As a result, its effect is usually discredited. For example, reminding people that the weather can influence their mood state eliminates the effect of the weather on judgments of life satisfaction (Schwarz and Clore 1983). Indeed, merely making the external source salient (e.g., by asking participants, "By the way, how is the weather out there?") seems to eliminate the effect of the extraneous cues on judgments (Schwarz and Clore 1983).

Given that, by definition, ambient scents are extraneous to the product, it can be expected that when they are made salient, their effect on product evaluations dissipates, especially when there is little or no semantic overlap with the product. However, when semantic overlap is high, it is expected that ambient scents will continue to affect evaluations, regardless of how salient they are. This reasoning leads to the following hypotheses:

> H_1: When the ambient scent is not salient, pleasant ambient scents increase product evaluations, regardless of their congruence with the product category.
>
> H_2: When the ambient scent is salient, pleasant ambient scents increase product evaluations only when they are congruent with the product category.

These hypotheses are the focus of Experiment 1.

Motivation to Correct

People usually assess the correctness of their initial reaction by elaborating on the differences and similarities between the extraneous cue and the target (Meyers-Levy and Sternthal 1993). This implies that consumers should be sufficiently motivated to elaborate on differences and similarities between the contextual cue and the target (Petty et al. 1993). In addition, correcting for erroneous influences is resource demanding and depends largely on the cognitive resources that a person is able or willing to devote to the judgmental task (Clore, Schwarz, and Conway 1994; Martin, Seta, and Crelia 1990). Thus, correcting for extraneous influences is more likely under conditions of high processing motivation. Formally,

> H_3: When processing motivation is low, pleasant ambient scents increase product evaluations, regardless of their congruence with the product category.
>
> H_4: When processing motivation is high, pleasant ambient scents increase product evaluations only when they are congruent with the product category.

These hypotheses are tested in Experiments 2 and 3. To increase the generalizability of the results, Experiment 2 investigates the moderating role of processing motivation by examining individual differences in consumers' need for cognition (NFC; Cacioppo and Petty 1982), and Experiment 3 focuses on differences in personal involvement with the evaluation task.

Experiment 1: Salience of Ambient Scents

Experiment 1 investigates whether consumers correct for the influence of pleasant ambient scents when they are made aware of their presence. As argued previously, it is expected that people are more likely to control for the influence of ambient scents when there is little or no semantic overlap between the scent and the product and when they are aware of the presence of an extraneous source (i.e., the ambient scent). As a result, when the ambient scent is made salient, consumers are more likely to correct for its influence when the scent is incongruent with the product category. When the extraneous scent is not salient, both congruent and incongruent pleasant scents are expected to increase product evaluations.

Design and Measures

The sample consisted of 80 undergraduate students, who received monetary compensation of $5 for their participation. A 2 (salience of ambient scent: salient versus not salient) × 2 (congruence of ambient scent: congruent versus not congruent) between-subjects design was used. In addition, a control condition was included in which no ambient scent was present. Participants were randomly assigned to one of the experimental conditions.

This experiment was the first in a battery of three (unrelated) consumer behavior experiments. All three experiments were conducted in different rooms (with a different experimenter in each room). Participants were received in groups of up to two and were asked to evaluate a new (fictitious) brand of orange juice (Sano), which was displayed in a color print advertisement that pictured the new brand, together with some oranges, and claimed, "Sano ... a natural source of vitamins." Attitudes toward the brand were assessed with two general evaluation items on nine-point scales (negative/positive and I dislike it/I like it; $r = .69$).

In the salient condition, consumers were told up front that an ambient scent was present in the environment (without mentioning the nature of the scent) and that the experiment was intended to investigate the extent to which environmental scents affect consumer behavior. In the nonsalient condition, there was no reference to the presence of the ambient scent.

In the scented conditions, the ambient scents were emitted by two pieces of scented papers (each containing three drops of essential oil) underneath the participants' desk. The scented papers were replaced before participants entered the room. The experiment was run in two identical rooms (i.e., similar furniture and same size), and only one scent per room per day was used to avoid mixing odors (the different experimental conditions were counterbalanced between rooms). The rooms were fully ventilated overnight. Each morning, the rooms were checked by at least two experimenters to exclude the possibility that scents remained in the room. In the congruent condition, a citrus scent was emitted, and in the noncongruent condition, a forest scent was used. A pretest ensured that both scents differed with respect to the congruency with the advertised product but did not differ with respect to their hedonic properties.

Pretest

Small glass bottles that contained a cotton ball with three to four drops of the essential oil were presented to 30 volunteer undergraduate students (adapted from Spangenberg, Crowley, and Henderson 1996). To avoid possible effects of color, opaque bottles were used. The bottles were labeled with random numbers and did not refer to the specific nature of the scents. Participants were asked to sniff the bottles at approximately six inches from the nose. They could sniff as many times as they considered necessary to answer the questions.

Congruency of the scents was assessed by asking participants about the appropriateness of the scent for the orange juice presented in the advertisement (see Morrin and Ratneshwar 2003). Answers were given on a nine-point scale ranging from "not at all appropriate" (1) to "highly appropriate" (9). The citrus scent was perceived as significantly more appropriate for the advertised product than the forest scent ($M_{citrus} = 4.77$ versus $M_{forest} = 1.57$; $t(29) = 7.59$, $p < .001$).

The hedonic properties of the scents were assessed with items from Fisher's (1974) environmental quality scale. Both scents were perceived as equally pleasurable (positive/negative, pleasurable/unpleasurable, comfortable/uncomfortable, good/bad, and attractive/unattractive; $\alpha = .93$) on a nine-point scale ($M_{citrus} = 4.33$ versus $M_{forest} = 3.79$; $t(29) = 1.60$, not significant [n.s.]). In addition, no differences were found in arousal between scent conditions (tense/relaxed, boring/stimulating, unlively/lively, and dull/bright; $\alpha = .76$; $M_{citrus} = 4.13$ versus $M_{forest} = 4.43$; $t(29) = 1.44$, n.s.).

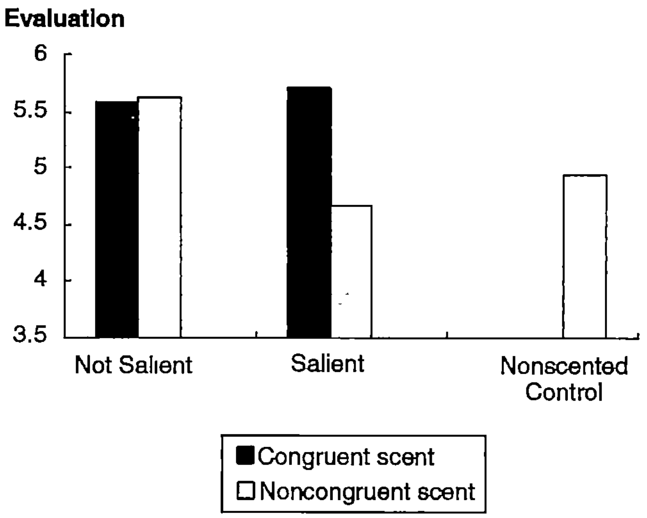
Results and Discussion

A two-way analysis of variance (ANOVA) was performed with salience and congruence of the ambient scent as independent variables and evaluation of the new brand of orange juice as the dependent variable. In addition, the effects of the congruent and the noncongruent scent were compared with the nonscented control condition for each salient condition separately. The results appear in Figure 1.

There was a main effect of congruence ($F(1, 60) = 4.80$, $p < .05$) and a marginally significant main effect of salience ($F(1, 60) = 3.34$, $p = .07$). More important, however, there was a significant interaction effect ($F(1, 60) = 5.87$, $p = .01$).

In support of H_1, when the ambient scent was not made salient, evaluations of the new brand of orange juice did not differ significantly as a function of congruence of the ambi-

FIGURE 1
Experiment 1: Mean Product Evaluations as a Function of Salience and Congruence of Ambient Scent



ent scent. Participants evaluated the brand equally positive when either a congruent or a noncongruent scent was present in the room ($M_{congruent} = 5.58$ versus $M_{not\ congruent} = 5.63$; $F(1, 60) < 1$). Additional analyses revealed that when the ambient scent was not made salient, evaluations in both the congruent and the noncongruent ambient scent conditions were significantly more positive than evaluations in the nonscented control condition ($M_{control} = 4.94$; $F(1, 51) = 5.44$, $p < .05$; and $F(1, 51) = 6.37$, $p < .05$, respectively).

Consistent with H_2, when the ambient scent was made salient, evaluations were more positive when the scent was congruent than when it was not congruent with the product category ($M_{congruent} = 5.71$ versus $M_{not\ congruent} = 4.67$; $F(1, 60) = 8.94$, $p < .01$). Additional analysis revealed that evaluations in the congruent condition were significantly more positive than those in the nonscented control condition ($F(1, 39) = 4.10$, $p = .05$). In contrast, no difference was found between evaluations in the noncongruent versus the control condition ($F(1, 39) < 1$).

In support of the predictions, consumers corrected for the influence of ambient incongruent scents only when the extraneous influence of the scents was made salient. In contrast, congruent ambient scents continued to influence product evaluations even when their extraneous influence became salient.

Whereas the influence of other affective cues (e.g., weather, music) is discounted as soon as their extraneous influence becomes salient (e.g., Pham 1998; Schwarz and Clore 1983), the influence of ambient scents seems to be stronger. They continue to affect product evaluations as long as they are congruent with the product category and perceived as appropriate for evaluations, regardless of whether the source of the extraneous influence becomes salient.

Experiment 2: Processing Motivation

The purpose of Experiment 2 is threefold. A first goal is to examine whether consumers control for the influence of pleasant ambient scents when they have sufficient processing motivation to correct for their extraneous influence. Filtering out extraneous influences requires that people have sufficient resources available to elaborate on differences and similarities between the contextual cue and the target (Petty et al. 1993). In addition, correcting for erroneous influences seems to be resource demanding because people carefully need to distinguish their reactions toward the target from their reactions toward the product (Martin 1986; Martin, Seta, and Crelia 1990). As a result, it can be expected that consumers correct for ambient scents only when they are sufficiently motivated to correct for possible extraneous influences. Consistent with the results obtained in Experiment 1 and with previous theories on the effects of congruent versus incongruent extraneous cues on evaluations (Meyers-Levy and Sternthal 1993), it is expected that this correction process is more likely for incongruent ambient scents.

A second purpose of Experiment 2 is to examine the mediating role of consumers' affective reactions in the relationship between ambient scents and product evaluations. As argued previously, it is assumed that ambient scents act like affective cues and that their effect on product evaluations is due to a misattribution of people's affective reactions (elicited by the scent) to their reaction to the product. Consequently, consumers' affective reactions should mediate the hypothesized effect of (in)congruent scents on product evaluations.

In Experiment 1, the ambient scent was emitted by attaching scented papers underneath participants' desks. This was done to make the ambient scent extraneous to the judgmental process and to avoid participants considering the ambient scent diagnostic for their evaluation by, for example, considering the scent a sensory sample of the advertised product ("the advertiser wants to give me some information about the smell and taste of the juice"). Although the extraneous manipulation is reasonable from a theoretical point of view, it weakens the external validity of the findings. Indeed, advertisers may be less inclined (and able) to scent the environments in which consumers read advertisements, but they may be more inclined to scent the advertisements themselves.

Therefore, a third objective of Experiment 2 is to generalize the results obtained in Experiment 1 to a more practical setting in which the advertisement, not the environment, is scented. Strictly speaking, when a scent surrounds an advertisement, it can be considered an extraneous cue because it is emitted by a source other than the product under evaluation.

Design and Procedure

One hundred eighteen undergraduate students volunteered to participate in this study. They were randomly assigned to a condition of a 2 (processing motivation: high versus low) × 3 (congruence of ambient scent: congruent, not congruent, or nonscented control) factorial between-subjects design.

As in Experiment 1, respondents participated in groups of no more than two and were asked to evaluate a new (fictitious) brand of grapefruit juice (Pompeltje) that was presented to them as a colored print advertisement. The advertisement displayed the new brand, together with the image of a grapefruit, and stated, "Pompeltje: All Fruit!" In addition, some product claims were provided that promoted the benefits of the juice.

In this context, NFC (Cacioppo and Petty 1982), a measured individual difference variable, was considered an index for consumers' processing motivation. People who score high on NFC usually think about things thoroughly and, in general, are motivated to elaborate on information. People who score low on the NFC measures are characterized by shallow information processing and, most of the time, are not motivated to process. Need for cognition (on a scale ranging from –3 to +3) was measured at the end of Experiment 2, after some unrelated filler tasks. Participants with mean scores below the NFC median ($Md_{NFC} = .97$) were assigned to the low-processing-motivation condition ($M_{low\ processing\ motivation} = .27$), and participants who scored greater than or equal to the median were assigned to the high-processing-motivation condition

($M_{\text{high processing motivation}}$ = 1.52). A t-test revealed that both conditions differed significantly with respect to their mean NFC scores (t(116) = 15.03, $p < .001$).

The advertisement was scented by attaching a scented paper (containing one to two drops of essential oil) to the flip side of the advertisement. In the congruent condition, the advertisement was scented with a citrus scent; and in the noncongruent condition, a lavender scent was used.

Pretest

Again, the congruence of the scents with the product advertised and their hedonic properties were assessed in a pretest. The procedure was similar to the one described in the pretest of Experiment 1. The citrus scent was perceived as significantly more appropriate for the advertised product than was the lavender scent (M_{citrus} = 4.06 versus M_{lavender} = 2.63; t(29) = 5.58, $p < .001$). In addition, both scents had similar hedonic qualities (pleasure measures: M_{citrus} = 4.33 versus M_{lavender} = 3.88; t(29) = 1.30, n.s.; arousal measures: M_{citrus} = 4.43 versus M_{lavender} = 4.17; t(29) = 1.18, n.s.).

Dependent Measures

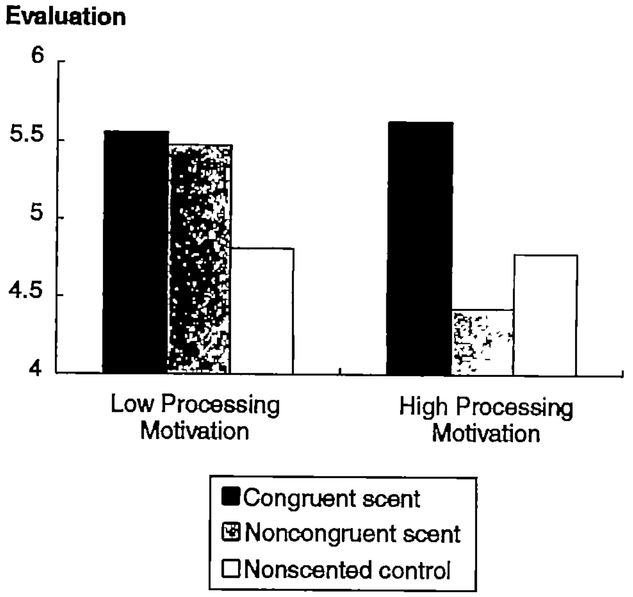
Evaluation of the new brand of grapefruit juice was assessed with four general evaluation items (negative/positive, undesirable/desirable, I dislike it/I like it, and bad/good; $\alpha = .90$). In addition, participants' affective and cognitive reactions were measured by asking them about their feelings (depressing/cheerful, ugly/beautiful, and irritating/nice; $\alpha = .74$) and cognitions (not useful/useful, not efficient/efficient, and informative/not informative; $\alpha = .76$) about the advertisement that was presented to them. A factor analysis revealed that the affect-related items loaded on the affective factor (explaining 33% of the total variance), whereas the cognition-related items loaded on the cognitive factor (explaining 36% of the total variance). All factor loadings were greater than .66. If ambient scents are indeed affective cues, congruence of ambient scents should mainly affect consumers' affective reactions (i.e., how pleasurable the scented advertisement is) and not, or to a lesser extent, consumers' cognitive reactions (i.e., how useful and interesting the scented advertisement is).

The ambient scent was not made salient in any of the conditions. Awareness of the scent was controlled for by asking participants to describe their experiences when participating in the experiment (because the experiment was designed presumably to examine people's experiences when participating in experiments in general) as precisely and as detailed as possible. They were also asked to "focus on the things you have noticed or that may have struck you." None of the participants spontaneously referred to the presence of the ambient scent.

Results and Discussion

Product evaluations. A two-way ANOVA revealed a main effect of processing motivation (F(1, 112) = 4.37, $p < .05$) and a main effect of congruence of ambient scent (F(2, 112) = 9.76, $p < .001$). More important, however, a significant interaction effect was observed (F(2, 112) = 5.00, $p < .01$). These results appear in Figure 2.

FIGURE 2
Experiment 2: Mean Product Evaluations as a Function of Processing Motivation and Congruence of Ambient Scent



Planned comparisons showed that in the low-processing-motivation condition, and consistent with H_3, evaluations in both the congruent ($M_{\text{congruent}}$ = 5.56) and the noncongruent ($M_{\text{not congruent}}$ = 5.47) condition were significantly more positive than those in the nonscented control condition (M_{control} = 4.81; F(1, 112) = 7.54, $p < .01$, and F(1, 112) = 5.40, $p < .05$, respectively). No difference was found between evaluations in the congruent and the noncongruent ambient scent conditions (F(1, 112) < 1).

In contrast, and consistent with H_4, when processing motivation was high, evaluations in the congruent condition ($M_{\text{congruent}}$ = 5.63) were significantly more positive than those in the noncongruent ($M_{\text{not congruent}}$ = 4.42) and nonscented control (M_{control} = 4.78; F(1, 112) = 18.78, $p < .001$, and F(1, 112) = 9.88, $p < .01$, respectively) conditions. No difference was found between evaluations in the noncongruent and nonscented control conditions (F(1, 112) = 1.81, $p > .18$).

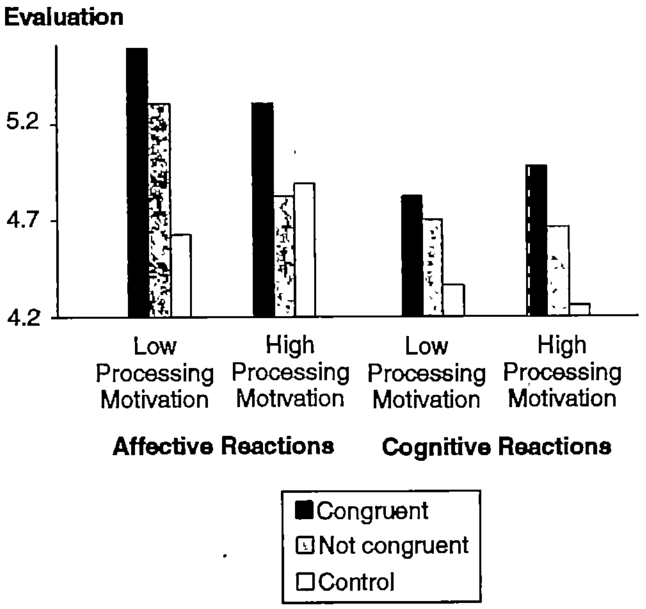
Affective versus cognitive reactions. A three-way ANOVA with processing motivation and congruence of ambient scent as between-subjects variables and reaction (affective versus cognitive) as a within-subjects variable was performed to gain insight into the effect of ambient scents on the specific nature of consumers' reactions. This analysis revealed a significant main effect of congruence (F(2, 112) = 7.73, $p < .001$) and a significant main effect of reaction (F(1, 112) = 42.90, $p < .001$). Overall, reactions were more positive when the ambient scent was congruent ($M_{\text{congruent}}$ = 5.17) than when it was not ($M_{\text{not congruent}}$ = 4.87) or when there was no scent present (M_{control} = 4.54). In addition, affective reactions were more positive than cognitive reactions ($M_{\text{affective}}$ = 5.08 versus $M_{\text{cognitive}}$ = 4.63).

More important, however, a significant three-way interaction was observed (F(2, 112) = 3.64, $p < .05$). The results appear in Figure 3.

Planned comparisons revealed a marginally significant two-way interaction between processing motivation and congruence on participants' affective reactions (F(2, 112) = 2.45, $p = .09$). When processing motivation was low, affective reactions were more positive when the ambient scent was either congruent ($M_{congruent} = 5.58$) or not congruent ($M_{not\ congruent} = 5.30$) compared with the nonscented control condition ($M_{control} = 4.63$; F(1, 112) = 15.37, $p < .001$, and F(1, 112) = 6.94, $p < .01$, respectively). When processing motivation was high, affective reactions were more positive than the nonscented control condition ($M_{control} = 4.89$) only when the scent was congruent ($M_{congruent} = 5.30$; F(1, 112) = 2.89, $p = .09$, marginally significant). No difference was found between affective reactions in the noncongruent condition ($M_{not\ congruent} = 4.82$) and in the nonscented control condition (F(1, 112) < 1). These results resemble those of participants' overall evaluation scores.

With regard to participants' cognitive reactions, only a significant main effect of congruence was found (F(1, 112) = 4.74, $p = .01$), indicating that compared with the nonscented control condition ($M_{control} = 4.31$), cognitive reactions were more positive when a congruent scent ($M_{congruent} = 4.90$) or a noncongruent scent ($M_{not\ congruent} = 4.68$) was present (F(1, 112) = 9.31, $p < .01$, and F(1, 112) = 3.56, $p = .06$, respectively). No difference in cognitive reactions was found between the congruent and the noncongruent conditions (F(1, 112) = 1.21, n.s.). These results suggest that the presence of a pleasant ambient scent, regardless of its level of congruence with the product category, is likely to increase consumers' cognitive reactions.

FIGURE 3
Experiment 2: Mean Affective and Cognitive Reactions as a Function of Processing Motivation and Congruence of Ambient Scent



Affective mediation. It is expected that consumers' affective reactions mediate the effect of the congruence of the ambient scent on product evaluations. Structural equation modeling (AMOS, implemented in SPSS) was used to investigate the mediating role of consumers' affective reactions in the influence of ambient scents on product evaluations.

As indicated by the results of the low-processing-motivation condition, no significant effect of congruent versus noncongruent scents was found on product evaluations (B = .08, n.s.). In the high-processing-motivation condition, the direct effect of the congruence of the ambient scent on product evaluations (B = .52, $p < .001$) was significantly mediated by consumers' affective reactions and not, or to a lesser extent, by their cognitive reactions. The results of the structural model ($\chi^2(40) = 67.04$, $p < .01$; comparative fit index = .87) indicate that when consumers' affective and cognitive reactions are included, the relationship between the congruence of the ambient scent and product evaluations drops to .33 ($p < .05$). Both consumers' affective (B = .56, $p < .01$) and cognitive (B = .33, $p < .05$) reactions are predictive for consumers' final evaluations, but congruence is a significant predictor only for consumers' affective reactions (B = .31, $p = .09$). Congruence does not predict consumers' cognitive reactions (B = .26, n.s.). Thus, as expected, consumers' affective reactions seem to mediate the relationship between the congruence of ambient scents and product evaluations.

As in Experiment 1, it was found that congruent extraneous scents continue to influence final product evaluation. In addition and consistent with expectations, consumers seem to control for the influence of incongruent scents only when they are sufficiently motivated to correct for erroneous influences. Moreover, consumers' affective reactions in response to the scent mediated the effect of congruence on product evaluations.

Experiment 3: Processing Motivation and Degree of Congruence

Experiments 1 and 2 show that the effects of pleasant ambient scents may be stronger than the effects of other affective cues in that as long as they are congruent with the product category, they continue to affect product evaluations even when the scent is made salient or when consumers have sufficient processing motivation to correct for extraneous influences. In addition, when a scent is incongruent, consumers seem to correct for it only when the scent becomes salient or when they have sufficient processing resources to correct for extraneous influences. These results can be explained by the notion that the stronger the semantic overlap between the ambient scent and the product, the more difficult it is to distinguish reactions toward extraneous cues from reactions toward the target. Indeed, the weaker the semantic link between the extraneous cue and the target, the fewer cognitive resources seem to be required to detect and filter out inappropriate extraneous influences (Meyers-Levy and Tybout 1997). As a result, it can be expected that the more incongruent the scent is with the product, the more likely it will be that consumers will correct for its extraneous influence. The main purpose of Experiment 3 is to investigate the effects of level of perceived congruence of the ambient scent on consumers' product evaluations. More specifically,

- H_5: When the ambient scent is highly congruent with the product category, it continues to affect product evaluations, regardless of consumers' processing motivation.
- H_6: When the ambient scent is of medium congruence with the product category, it affects product evaluations when processing motivation is low but not when processing motivation is high.
- H_7: When the ambient scent is incongruent with the product category, it does not affect product evaluations

Thus, consistent with the results of Experiments 1 and 2, scents that are highly congruent with the product category are expected to continue to affect judgments, regardless of whether consumers are sufficiently motivated to correct for extraneous influences. In addition, however, it is argued that whether consumers correct for the effects of incongruent ambient scents depends on the degree of incongruence between the scent and the product. Given that fewer resources seem to be required to detect and filter out highly incongruent external influences, it is proposed that whereas medium incongruent scents are corrected for only when consumers are sufficiently motivated to discount extraneous influences, the effects of highly incongruent scents are corrected for regardless of consumers' processing motivations.

Another purpose of Experiment 3 is to generalize the results obtained in Experiments 1 and 2 to a retail environment. Whereas participants in Experiments 1 and 2 were asked to evaluate new brands in an advertisement context, Experiment 3 investigates the effect of ambient scents in a setting in which people are asked to evaluate authentic products. In addition, instead of emitting scents with scented papers (either in the environment or attached on the flip side of an advertisement), the room was scented with an electric diffuser, as in in-store environments.

Design and Procedure

The sample consisted of 75 undergraduate students, who were compensated $5 for their participation. The design was a 2 (processing motivation: high versus low) × 2 (presence of ambient scent: pleasant banana scent versus nonscented control) × 3 (product: banana, apple, or tomato) mixed-subjects design.

As in Experiments 1 and 2, groups comprised no more than two participants, and they were asked to evaluate several products that were highly congruent, of medium congruence, or incongruent with the ambient scent that was present in the environment. In the scented conditions, the smell of bananas was distributed by means of an electric diffuser that was placed out of participants' sight. The experiment was the first in a battery of three (unrelated) consumer behavior experiments that were all conducted in different rooms. Each morning, the consumer lab was scented ten minutes before participants were scheduled to arrive. The diffuser emitted the scent continuously into the lab's atmosphere at a medium intensity level. The lab was fully ventilated overnight.

The study was presented as being about the evaluation of "the fruits and vegetables of a new (fictitious) fruit auction company, 'Westham Corporation.'" Processing motivation, a between-subjects factor, was manipulated by increasing participants' personal involvement with the task (see also MacInnis and Park 1991). In the low-processing-motivation condition, participants were instructed to "evaluate these fruits and vegetables that are auctioned by the new fruit auction company Westham. This company is conducting a lot of interviews these days." In the high-processing-motivation condition, participants were instructed to "evaluate these fruits and vegetables that are auctioned by the new fruit auction company Westham. They appreciate your opinion a lot, and after your evaluation task, a manager of the corporation will interview you about your opinion." Afterward, participants were asked to evaluate three products: a banana (highly congruent with the ambient scent), an apple (medium congruent with the ambient scent), and a tomato (incongruent with the ambient scent). The experimenter handed one fruit product at a time to participants, and they were asked to evaluate them subsequently. The order in which the fruits were presented to participants was counterbalanced across conditions. A pretest revealed that participants perceived the banana scent as more appropriate for a banana ($M_{banana} = 4.43$) than for an apple ($M_{apple} = 3.57$; $t(29) = 10.47$, $p < .01$) or for a tomato ($M_{tomato} = 1.73$; $t(29) = 30.16$, $p < .001$). In addition, the banana scent was more appropriate for an apple than for a tomato ($t(29) = 7.73$, $p < .01$).

Dependent Measures

Evaluations of the fruits and vegetables were assessed by means of four nine-point evaluation items (i.e., fresh, good, I like it, and tasty; $\alpha = .52$). In addition, the effectiveness of the processing motivation manipulation was assessed by asking participants about their involvement when evaluating the products (this research is unimportant/important to me, does not concern/concerns me, means nothing to me/means a lot to me; $\alpha = .81$). Again, no reference was made to the presence of the ambient scent. In the debriefing, none of the participants indicated that they noticed the presence of a scent.

Results and Discussion

Manipulation check processing motivation. A two-way ANOVA with processing motivation and the presence of an ambient scent as between-subjects factors and involvement with the task as a dependent variable revealed only a main effect of processing motivation ($F(1, 71) = 5.23$, $p < .05$). Participants in the low-processing-motivation condition reported being less involved ($M_{\text{low processing motivation}} = 3.58$) than participants in the high-processing-motivation condition ($M_{\text{high processing motivation}} = 4.24$).

Product evaluations. The ANOVA with processing motivation and the presence of an ambient scent as between-subjects variables and product as the within-subjects variable revealed a significant three-way interaction effect ($F(2, 142) = 3.31$, $p < .05$). Other than a marginally significant main effect of the presence of an ambient scent ($F(1, 71) = 3.63$, $p = .06$), indicating that, overall, evaluations were more positive when a scent was present than when a scent was not present ($M_{scent} = 2.06$ versus $M_{control} = 1.16$), no other effects were significant. The results appear in Figure 4.

When the product was highly congruent with the scent, neither the main effect of processing motivation nor the interaction effect was significant (both $Fs < 1$). Consistent with H_5, a significant main effect of the presence of an ambient scent was found ($F(1, 71) = 4.97$, $p < .05$), indicating that evaluations of the banana were more positive when a scent was present than when a scent was not present ($M_{banana\ scent} = 5.48$ versus $M_{banana\ control} = 4.98$), regardless of processing motivation.

When the product was of medium congruence with the scent, the interaction between the presence of an ambient scent and processing motivation was significant ($F(1, 71) = 5.64$, $p < .05$). Consistent with H_6 and with the results of Experiments 1 and 2, when processing motivation was low, evaluations of the apple were more positive when a scent was present than when a scent was not present ($M_{apple\ scent} = 5.69$ versus $M_{apple\ control} = 4.89$; $F(1, 71) = 5.67$, $p < .05$). When processing motivation was high, no difference was found between the scented condition and the nonscented control condition ($F(1, 71) = 1.05$, n.s.; $M_{apple\ scent} = 4.94$ versus $M_{apple\ control} = 5.29$). Other than this interaction effect, no other effects were significant (all $Fs < 1$).

Finally, when the product was highly incongruent with the scent, none of the main effects (both $Fs < 1$) nor the interaction effect ($F(1, 71) = 1.29$, n.s.) was significant. Consistent with H_7, when the scent was highly incongruent, consumers found it easier to control for its possible erroneous influence, regardless of their processing motivation.

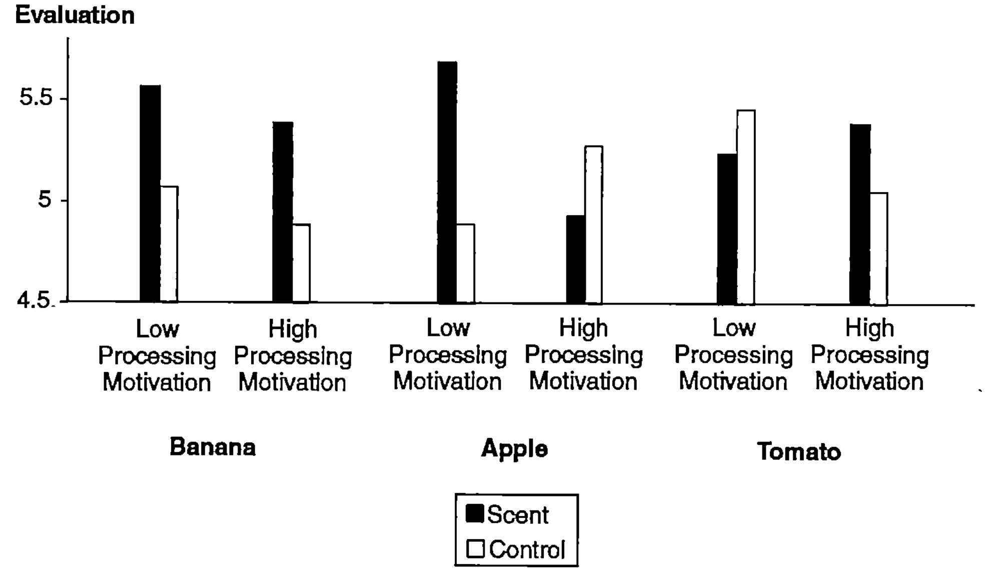
These results support H_5–H_7 and demonstrate that whether consumers correct for the influence of ambient scents depends on the level of congruence of the scents with the product category. The stronger the semantic linkage between the scent and the product, the more difficult it is to correct for extraneous influences, and the more resources are needed to filter out the extraneous influences. Consistent with this, when the product was highly incongruent with the ambient scents, respondents corrected for extraneous influences, even when they were not motivated to process. In contrast, when overlap between the scent and the product was high, scents continued to influence participants' evaluations, regardless of processing motivation. Only when the product was of medium congruence with the scent did consumers correct for its influence when they were highly motivated to process.

General Discussion

The three experiments demonstrate that ambient scents, which are extraneous to consumers' judgmental process, can have a strong influence on consumers' product evaluations. As long as ambient scents are congruent with the product category, they continue to affect product judgments, even when they become salient or when consumers are sufficiently motivated to discount their potential influences. In addition, scents that are not congruent with the product category can still influence consumers' product evaluations

FIGURE 4

Experiment 3: Mean Product Evaluations as a Function of Processing Motivation, Presence of the Ambient Scent, and Product



when they are not salient or when consumers lack processing motivation. Only when the ambient scent is completely incongruent with the product do consumers seem to be able to cope well with their possible influence.

This research reveals three important conditions on which scent effects are contingent: (1) the congruence of the scent with the product category, (2) the salience of the extraneous influence, and (3) consumers' motivation to correct for possible extraneous influences. Whereas previous studies have convincingly shown that congruent scents have beneficial effects on variety seeking, information search, and recall (e.g., Mitchell, Kahn, and Knasko 1995; Morrin and Ratneshwar 2003), research has been less conclusive about the impact of congruency on product evaluations (Knasko 1995; Morrin and Ratneshwar 2003). By taking into account not only the congruence of a scent with the product category but also its salience and consumers' processing motivation, the current findings clarify the mixed results that were previously obtained with regard to the impact of ambient scents on product evaluations and the moderating role of consumers' awareness of ambient scents.

The results further show that ambient scents can act as affective cues, providing affective information for product evaluation. The results of Experiment 2 suggest that consumers' affective reactions (as a result of the pleasantness of the scent) are predictive for product evaluation and mediate the effect of congruence of the scent on judgments. However, note that the results do not exclude the possibility that ambient scents influence cognitive reactions, as, for example, Chebat and Michon (2003) suggest. Indeed, the results obtained in Experiment 2 suggest that cognitive reactions were more positive as a result of the presence of a pleasant ambient scent. However, cognitive reactions were affected regardless of the congruence of the scent and consumers' processing motivation, and they did not mediate the effect of congruence on product evaluations. Further research is necessary to investigate these unexpected results in closer detail and to clarify how the presence of a pleasant ambient scent influences consumers' cognitive reactions.

Research related to the impact of affective cues on persuasion has mostly concentrated on the pleasantness dimension of the affective cues (i.e., how positive or negative the cue is). In addition to their level of pleasantness, ambient scents provide more semantic information that increases or decreases their congruence with the target product. According to the results obtained in Experiments 1–3, this congruence dimension is a necessary condition for explaining the effects of ambient scents on product evaluations. Whether congruence is also important for explaining the effects of other affective cues remains unclear, but given that music, for example, can also provide information that goes beyond its level of pleasantness (North, Hargreaves, and McKendrick 1999), it would be interesting for further research to investigate how important the congruence dimension is for affective cues other than ambient scents.

When participants discounted the influence of ambient scents, they seemed to be able to filter out the extraneous influences. That is, contrast or overcorrection effects (resulting in more negative evaluations), which are frequently observed in the contextual correction literature (e.g., Martin, Seta, and Crelia 1990), were never observed. These results suggest that the use of pleasant ambient scents never hurts. However, further research is necessary to investigate the possible occurrence of overcorrection in the presence of incongruent ambient scents. In addition, whereas it is assumed in general that correction for extraneous influences is a conscious and deliberative process, the results of the thought protocols suggest that correction for ambient scents operates below consumers' level of awareness. Indeed, participants did not refer spontaneously to the (influence of the) ambient scent in any of the thought protocols. Further research is needed to investigate the degree to which consumers are truly conscious about this correction process.

Managerial Implications

Ambient scents can be used in a variety of practical settings. Retailers can use in-store scent diffusers, advertisers can use scented advertisements, package designers can use scented wrappings, and so forth. The results of this study suggest that as long as the ambient scent is congruent with the product category, it will have a positive effect on consumers' product evaluations, regardless of how salient the scent is or how motivated (or involved) consumers are. This implies that using ambient scents to embellish store atmospherics or to increase consumers' overall evaluation may be more effective than using other affective cues, such as music or colors. Whereas the effects of other ambient stimuli are believed to dissipate when these cues become salient or when consumers' processing motivation increases, the effects of congruent ambient scents continue to affect evaluations.

Retailers may be worried about the effects of incongruent scents (e.g., when the scent of the bakery department reaches the detergents department, or vice versa; when the Channel No. 5 fragrance emitted at the entrance reaches the fresh vegetables and fruit department). Whether these scents influence consumers' product evaluations largely depends on how salient the scent is or how motivated consumers are. As long as the ambient scent is not too salient or consumers' processing motivation (as a result of either individual differences or increased purchase/product involvement) is not too high, these pleasant scents will have a positive impact on product evaluations. More important, however, the results suggest that these incongruent scents will not hurt evaluations; evaluations were not significantly worse in any of the incongruent scent conditions than in the nonscented control conditions.

In summary, using pleasant smells can create promising opportunities for marketing practice. On the one hand, ambient scents are more effective as their congruence with the product category increases. On the other hand, marketers should not be worried about effects of these scents on incongruent product categories, because incongruent scents do not seem to hurt overall product evaluations.

REFERENCES

Aggleton, John P. and Mortimer Mishkin (1986), "The Amygdala: Sensory Gateway to the Emotions," in *Emotion: Theory, Research and Experience*: *Biological Foundations of Emotions*, Vol. 3, Robert Plutchik and Henry Kellerman, eds. Orlando, FL: Academic Press, 281–99.

Baron, Robert A. (1990), "Environmentally Induced Positive Affect: Its Impact on Self-Efficacy, Task Performance, Negotiation, and Conflict," *Journal of Applied Social Psychology*, 20 (March), 368–84.

——— and Marna I. Bronfen (1994), "A Whiff of Reality: Empirical Evidence Concerning the Effects of Pleasant Fragrances on Work-Related Behavior," *Journal of Applied Social Psychology*, 24 (July), 1179-1203.

——— and Jill Thomley (1994), "A Whiff of Reality: Positive Affect as a Potential Mediator of the Effects of Pleasant Fragrances on Task Performance and Helping," *Environment and Behavior*, 26 (November), 766–84.

Bitner, Mary Jo (1992), "Servicescapes: The Impact of Physical Surroundings on Customers and Employees," *Journal of Marketing*, 56 (April), 57–71.

Bone, Paula Fitzgerald and Pam Scholder Ellen (1999), "Scents in the Marketplace: Explaining a Fraction of Olfaction," *Journal of Retailing*, 75 (Summer), 243–62.

——— and Swati Jantrania (1992), "Olfaction as a Cue for Product Quality," *Marketing Letters*, 3 (July), 289–96.

Bosmans, Anick and Hans Baumgartner (2005), "Goal-Relevant Emotional Information: When Extraneous Affect Leads to Persuasion and When It Does Not," *Journal of Consumer Research*, 32 (December), 424–34

Cacioppo, John T. and Richard E Petty (1982), "The Need for Cognition," *Journal of Personality and Social Psychology*, 42 (January), 116–31.

Campbell, Margaret C. and Amna Kirmani (2000), "Consumers' Use of Persuasion Knowledge: The Effects of Accessibility and Cognitive Capacity on Perceptions of an Influence Agent," *Journal of Consumer Research*, 27 (June), 69–83.

Cann, Arnie and Debra A. Ross (1989), "Olfactory Stimuli as Context Cues in Human Memory," *American Journal of Psychology*, 102 (Spring), 91–102.

Chebat, Jean-Charles and Richard Michon (2003), "Impact of Ambient Odors on Mall Shoppers' Emotions, Cognition, and Spending: A Test of Competitive Causal Theories," *Journal of Business Research*, 56 (July), 529–39.

Clore, Gerald F., Norbert Schwarz, and Michael Conway (1994), "Affective Causes and Consequences of Social Information Processing," in *Handbook of Social Cognition*, 2d ed., Robert S. Wyer and Thomas K. Srull, eds. Tuxedo Park, NY: Lawrence Erlbaum Associates, 323–417.

Ehrlichman, Howard and Jack N. Halpern (1988), "Affect and Memory. Effects of Pleasant and Unpleasant Odors on Retrieval of Happy and Unhappy Memories," *Journal of Personality and Social Psychology*, 55 (November), 769–79.

Ellison, Sarah and E White (2000), "When There's More to an Ad Than Meets the Eye," *The Wall Street Journal*, (November 24), A11.

Engen, Trygg (1982), *The Perception of Odors*. New York: Academic Press.

Fisher, Jeffrey D. (1974), "Situation-Specific Variables as Determinants of Perceived Environmental Aesthetic Quality and Perceived Crowdedness," *Journal of Research in Personality*, 8 (August), 177–88.

Gorn, Gerald, Michel Tuan Pham, and Leo Yatming Sin (2001), "When Arousal Influences Ad Evaluation and Valence Does Not (and Vice Versa)," *Journal of Consumer Psychology*, 11 (July), 43–55.

Herz, Rachel S. and Trygg Engen (1996), "Odor Memory: Review and Analysis," *Psychonomic Bulletin & Review*, 3 (September), 300–313.

Hirsch, Alan (1995), "Effects of Ambient Odors on Slot Machine Usage in a Las Vegas Casino," *Psychology and Marketing*, 12 (October), 585–94.

Knasko, Susan C. (1989), "Ambient Odor and Shopping Behavior," *Chemical Senses*, 14 (94), 718.

——— (1995), "Pleasant Odors and Congruency: Effects on Approach Behavior," *Chemical Senses*, 20 (5), 479–87.

MacInnis, Deborah J. and C. Whan Park (1991), "The Differential Role of Characteristics of Music on High- and Low-Involvement Consumers' Processing of Ads," *Journal of Consumer Research*, 19 (September), 161–73.

Martin, Leonard L. (1986), "Set/Reset: Use and Disuse of Conception Impression Formation," *Journal of Personality and Social Psychology*, 51 (September), 493–504.

———, John J. Seta, and Rick A. Crelia (1990), "Assimilation and Contrast as a Function of People's Willingness and Ability to Expend Effort in Forming an Impression," *Journal of Personality and Social Psychology*, 59 (July), 27–37.

Meyers-Levy, Joan and Brian Sternthal (1993), "A Two-Factor Explanation of Assimilation and Contrast Effects," *Journal of Marketing Research*, 30 (August), 359–68.

——— and Alice M. Tybout (1997), "Context Effects at Encoding and Judgment in Consumption Settings: The Role of Cognitive Resources," *Journal of Consumer Research*, 24 (June), 1–14.

Mitchell, Deborah J., Barbara E. Kahn, and Susan C. Knasko (1995), "There's Something in the Air: Effects of Congruent or Incongruent Ambient Odor on Consumer Decision Making," *Journal of Consumer Research*, 22 (September), 229–38.

Morrin, Maureen and S. Ratneshwar (2003), "Does It Make Sense to Use Scents to Enhance Brand Memory?" *Journal of Marketing Research*, 40 (February), 10–25.

North, Adrian C., David J. Hargreaves, and Jennifer McKendrick (1999), "The Influence of In-Store Music on Wine Selections," *Journal of Applied Psychology*, 84 (April), 271–76.

Petty, Richard E., David W. Schumann, Steven A. Richman, and Alan J Strathman (1993), "Positive Mood and Persuasion: Different Roles for Affect Under High- and Low-Elaboration Conditions," *Journal of Personality and Social Psychology*, 64 (January), 5–20.

Pham, Michel Tuan (1998), "Representativeness, Relevance, and the Use of Feelings in Decision Making," *Journal of Consumer Research*, 25 (September), 144–59.

Raghunathan, Rajagopal, Michel Tuan Pham, and Kim Corfman (2006), "Informational Properties of Anxiety and Sadness, and Displaced Coping," *Journal of Consumer Research*, 32 (March), 596–602.

Rotton, James, Timothy Barry, James Frey, and Edgardo Soler (1978), "Air Pollution and Interpersonal Attraction," *Journal of Applied Social Psychology*, 8 (January–March), 57–71.

Schwarz, Norbert (1990), "Feelings as Information: Informational and Motivational Functions of Affective States," in *Handbook of Motivation and Cognition*, Vol. 2, E. Tory Higgins and Richard M. Sorrentino, eds. New York: Guilford Press, 527–61.

——— and Herbert Bless (1992), "Constructing Reality and Its Alternatives: An Inclusion/Exclusion Model of Assimilation and Contrast Effects in Social Judgment," in *The Construction of Social Judgments*, Leonard L Martin and Abraham Tesser, eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 217–45.

——— and Gerald L. Clore (1983), "Mood, Misattribution, and Judgments of Well-Being: Informative and Directive Functions of Affective States," *Journal of Personality and Social Psychology*, 45 (September), 513–23.

Spangenberg, Eric R., Ayn E. Crowley, and Pamela W. Henderson (1996), "Improving the Store Environment: Do Olfactory Cues Affect Evaluations and Behavior?" *Journal of Marketing*, 60 (April), 67–80.

Wegener, Duane T. and Richard E. Petty (2001), "Understanding Effects of Mood Through the Elaboration Likelihood and Flexible Correction Models," in *Theories of Mood and Cognition*, Leonard L. Martin and Gerald L Clore, eds Mahwaw, NJ: Lawrence Erlbaum Associates, 177–210.

Zeithaml, Valarie A. and Mary Jo Bitner (2000), *Services Marketing: Integrating Customer Focus Across the Firm*, 2d ed. New York: McGraw-Hill.

WAYNE STATE UNIVERSITY

DEAN

Wayne State University is one of the country's leading urban doctoral-extensive universities, providing access to a world-class education at a great value. Wayne State's 11 schools and colleges offer more than 350 major subject areas to over 33,000 graduate and undergraduate students. Located in the heart of Detroit's University Cultural Center, the home of renowned museums, galleries and theatres, the WSU main campus encompasses 203 acres of beautifully landscaped walkways and gathering spots, linking 100 education and research buildings. A number of new residential communities and a new 75-acres research and technology park contribute to the vitality of campus life.

With 58 full-time faculty, the AACSB International Accredited School of Business Administration serves nearly 2500 undergraduate and over 1200 graduate students pursuing degrees in the fields of accounting, finance, information systems & manufacturing, logistics, management, and marketing Currently, the School is in the final stages of approval of a doctoral program. The MBA program, which includes one of the nation's largest part-time programs is supported by an award-winning online education initiative Preparing business leaders for more than 60 years, the School's 30,000 alumni include many who hold key executive positions in major corporations, nonprofit organizations, and government agencies

The University is seeking a Dean who can lead the School in a mission of academic excellence. The Dean should be an experienced academic administrator or business executive who can lead the faculty and stakeholders in furthering the undergraduate and graduate business degree programs and faculty research and development. The Dean is responsible for advancing the School's strategy and leveraging the School's resources to build its brand identity, alumni loyalty, and external relations The Dean will be expected to identify and build on opportunities to work with business firms located throughout the state of Michigan and around the globe. Wayne State University, in the midst of its first-ever capital campaign, is raising $500 million for university-wide initiatives that will transform the learning and teaching experience on campus. The Dean will work with University officials, alumni, leaders of Fortune 500 companies, and prominent regional business firms to raise the $45 million targeted for construction of a new building for the School, and endowments for student scholarships and chaired professorships. The Dean should have a demonstrated record of accomplishment in a collegial leadership position and possess excellent interpersonal and communication skills. Experience with AACSB International accreditation is preferred

The Dean reports to the Provost and Senior Vice President for Academic Affairs The salary will be competitive and commensurate with experience and qualifications Review of applications will begin immediately and continue until the position is filled Applications must include a letter of intent, curriculum vitae, the names, addresses, and telephone numbers of three references, any evidence of recent scholarly activity, and evidence of other accomplishments Electronic applications must be sent to **https://jobs.wayne.edu** Inquiries should be directed to: **Dean Ralph Kummler, Search Chair, Wayne State University, College of Engineering, 1150 Engineering Building, Detroit, MI 48202; Phone: 313-577-3775; Fax: 313-577-5300; email: rkummler@eng.wayne.edu**

Please refer to posting number 032770.

Wayne State University is an Equal Opportunity/Affirmative Action Employer.

Stacy L. Wood & C. Page Moreau

From Fear to Loathing? How Emotion Influences the Evaluation and Early Use of Innovations

Innovation adoption is rarely a short process for consumers; accordingly, recent research has explored adoption as a dynamic process that is characterized by changing patterns, or diffusion, of consumer use of the innovation. This research suggests that adoption is rarely a *neutral* process and that consumers can experience strong emotions in the initial use of innovations. However, given such emotions, two opposing arguments can be made as to whether the inclusion of emotional responses increases the predictive power of traditional models of diffusion. On the one hand, experienced emotion may simply be a function of gained benefits and, as such, may already be captured in extant models through cognitive assessments of net benefits. On the other hand, and as data from two empirical and longitudinal studies demonstrate, the learning process is potentially emotion generating (independent of net benefits), and this emotion colors product evaluations. The emotional influence is sizable and, importantly, not a straightforward case of "easier is better." In this work, the authors present the E^3 (expectation → emotion → evaluation) model, which describes how managers can better predict and influence the successful diffusion of complex technological products.

As technological innovations reach the consumer market at an unprecedented pace, recent reports have highlighted the influence of consumers' expectations of and actual experiences in using a product on a product's ultimate success. For example, perceptions of usage difficulty have caused a significant number of consumers to delay purchases (e.g., 48% of potential digital camera buyers), and actual usage difficulty has caused them to return previously purchased products (e.g., 30% of all home networking products; Endt 2004, p. E8). Epitomizing a company caught by these trends, Royal Philips Electronics, which has suffered as a result of consumers' rejection of complex new products, recorded huge losses in 2000 and 2001 and reduced its workforce by 55,000 employees (*The Economist* 2004). To understand the scope of the problem, Chief Executive Officer Gerard Kleisterlee sent 100 Philips managers home one weekend with an assignment to get various Philips gadgets to operate. Many failed and "returned frustrated and some even angry," whereas "another group that succeeded returned quite proud" (Endt 2004, p. E8). These managers were struck by the strength of their own response to product complexity in use, but we note the emotional nature of the managers' responses.

As this example and our own experiences show, learning to use a new product can evoke an emotional response, independent of the emotions produced by the attributes or benefits of the product itself. The goal of this research is to understand the underlying causes of this type of emotional response and study its ramifications for complex product innovations.[1] Although the majority of prior innovation research has focused on a point of adoption, the field has observed recent calls for more process-oriented research (e.g., Golder and Tellis 1998; Rogers 1996) to look beyond purchase and address how early use affects ultimate adoption (Shih and Venkatesh 2004). This usage focus is increasingly important as managers strive to understand why some consumers fail to keep, or to keep using, new products. To better understand this consumer-centric process of innovation, we develop the E^3 (expectation → emotion → evaluation) model of emotional influence, test the model in two studies, and describe how managers can use the findings to help predict and influence new product success.

Stacy L. Wood is Associate Professor of Marketing, Moore School of Business, University of South Carolina (e-mail: wood@moore.sc.edu). C. Page Moreau is Assistant Professor of Marketing, Leeds School of Business, University of Colorado, Boulder (e-mail: page.moreau@colorado.edu). The authors thank Bill Bearden, Meg Campbell, Bill Dillon, Paul Herr, Amna Kirmani, Don Lehmann, and John Lynch for their comments and assistance in all phases of this research, and they thank Jen Dale for her research assistance. They also gratefully acknowledge research funding from the Leeds School of Business and the Moore School of Business and the insightful comments of the three anonymous *JM* reviewers. Both authors contributed equally to this work.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

---

[1] Although many of the 4000 diffusion articles and books published to date have demonstrated that perceived product complexity causes consumers to delay or avoid the purchase of a new product, fewer than 1% have examined the influence of perceived complexity on postadoption behavior (Rogers 1996, p. 91). This sparse attention is surprising, given that consumers' initial usage experiences are complex, especially for technological products (Mick and Fournier 1998), and are likely to influence product return rate, future brand loyalty, word-of-mouth communications, and desire to purchase other technologies.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Expectation, Emotion, and Evaluation

It has been well established that new product success can be reliably predicted by an evaluation of the product's benefits, costs, and relative advantage over competing alternatives (Moreau, Lehmann, and Markman 2001; Rogers 1996). In Rogers's (1996) classic work on diffusion, he also identifies the innovation's "complexity" (i.e., the perceived magnitude of the learning costs required to achieve its benefits) as an indicator of an innovation's success. Complexity is typically indicative of slower diffusion rates (Rogers 1996) and can even create disutility through "feature fatigue" (Thompson, Hamilton, and Rust 2005). In the formative period of trial and early use of new, complex products, consumers learn how to achieve promised benefits. Rogers calls this "how-to" knowledge and acknowledges that it has been overlooked to date as a key factor in diffusion success. However, interest in use is growing; to wit, Shih and Venkatesh (2004) recently identified "usage diffusion" as an important conceptualization to understand better when consumers will fully integrate a newly adopted innovation into their lives and thus create a strong foundation for the innovation's current profitability and future growth.

Simply bringing consumers to the point of trying a new innovation, let alone establishing regular use, is a difficult process (Meuter et al. 2005). Thus, if firms expend great effort to bring consumers to the point of trial, it seems important that firms then act to encourage a trial environment that is conducive to positive evaluation. Our research demonstrates how consumers' emotions play an influential role in early usage environments. We show that a significant proportion of these emotions results from the disconfirmation of consumer expectations about ease of use, implying that a firm may either predict or influence the consumer's emotional experience. Perhaps the most significant contribution of this research is the demonstration that emotions experienced in early use of the new product are not simply reflective of achieved benefits (or risks) but result directly from learning efforts. These emotions, both immediately and over time, significantly influence product evaluations, even after we account for the traditional diffusion indicators.

The E³ Model of Emotional Influence

We present a five-step conceptual model that describes how emotion influences the evaluation of complex innovations. This model fits well within the traditional models of diffusion but is unique in its focus on the important how-to period when consumers first learn to use a new product. The five-step E³ model appears in Figure 1. The core steps that describe consumers' experience of early use and the theoretical process by which emotions may influence diffusion (Steps 2, 3, and 4) are foundational to our theoretical contribution. The model also addresses how firms can either predict or influence the process through relevant and actionable antecedents (Step 1) and whether relatively ephemeral emotions can affect more lasting evaluations (Step 5).

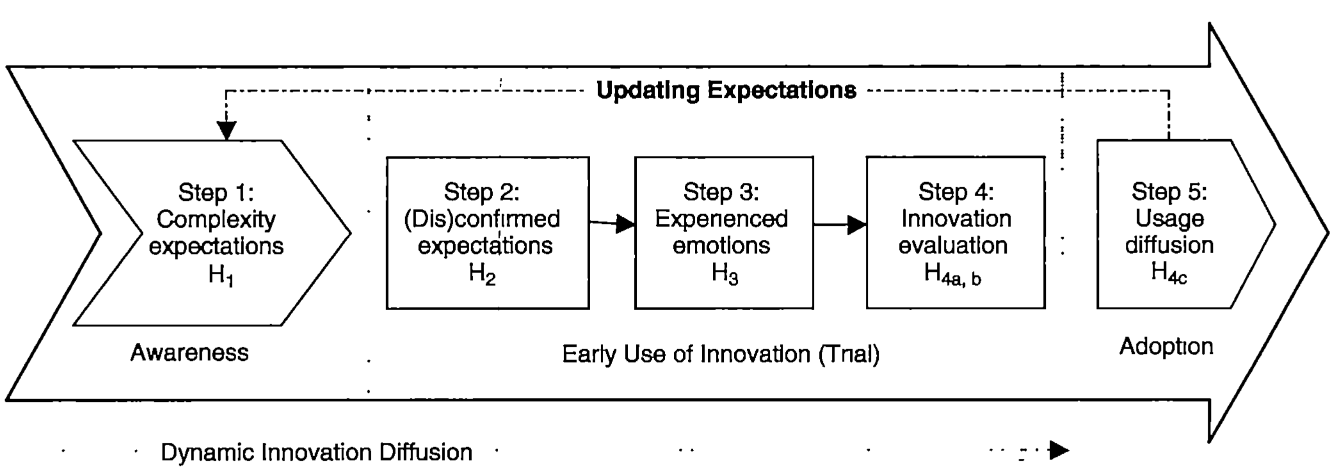
Steps 1 and 2: Complexity Expectations and Their (Dis)confirmation

In general, expectations matter. In marketing, the power of expectations is demonstrated in satisfaction research (e.g., Oliver and Swan 1989); thus, we consider the types of expectations that may be particularly influential in the early stages of new product learning. We posit that complexity expectations (e.g., "How difficult will it be to begin using this?") are critical for predicting both emotions and product evaluations.

Defining complexity expectations. We define "complexity expectations" as a priori predictions about ease of use, learning time, and learning difficulty. Theoretically, and akin to gap models of satisfaction, this conceptualization implies that a new product experience may be challenging but not necessarily disappointing if difficulty is expected.

This raises two issues that require careful conceptual distinction. First, although the satisfaction literature emphasizes the importance of expectations, these are expectations

FIGURE 1
The E³ Model of the Influence of Complexity Expectations on Innovation Evaluation and Diffusion


Updating Expectations
Step 1: Complexity expectations H_1
Step 2: (Dis)confirmed expectations H_2
Step 3: Experienced emotions H_3
Step 4: Innovation evaluation $H_{4a,b}$
Step 5: Usage diffusion H_{4c}
Awareness
Early Use of Innovation (Trial)
Adoption
Dynamic Innovation Diffusion

for the products' promised net benefits (e.g., Cadotte, Woodruff, and Jenkins 1987) and thus are captured in existing diffusion models through Rogers's (1996) long-standing inclusion of net benefits. This work is not a mere replication of previous satisfaction research. Here, we address consumers' expectations not for the product's ultimate benefits but for its complexity in learning. In other words, we focus on what consumers expect during their early usage experience (e.g., the rate at which they can achieve those benefits, the ease of learning, and their own affective response), and we posit that this is not a trivial construct. Thus, we investigate the potential for complexity expectations to influence product evaluation beyond the strong influence of the expected (and perceived) benefits and costs of the product. As we noted previously, extant diffusion models already successfully predict responses to new products (e.g., Moreau, Lehmann, and Markman 2001; Olshavsky and Spreng 1996), and the demonstration of any additive influence due to consumers' emotional response to initial usage complexity would have broad ramifications for new product research.

Second, complexity expectations should be more impactful when products are innovative than when they are new additions to well-understood product categories.[2] For some new products (e.g., a new orange juice), there is little uncertainty about the difficulty of use (Phillips and Baumgartner 2002). However, in cases in which the product requires adaptation or learning (e.g., Palm Pilots), usage uncertainty abounds (Oliver and Winer 1987). Thus, complexity expectations are likely to be influential only if products are both new and relatively complex. This leads to what is perhaps a marketplace paradox: The emotional experience of initial use is more influential for technological or functional innovations (e.g., a computer program, a Global Positioning System in an automobile) than for simple experiential or aesthetic products (e.g., a book, a music CD, a candle), which are typically associated with (and designed to elicit) an emotional response. However, this paradox is not so counterintuitive in light of the notion that the emotional response we investigate herein is that which arises from learning, not that which is tied directly to product attributes or benefits.

How are complexity expectations formed? For firms to be able to predict or influence consumers' complexity expectations, they must understand how such expectations are formed. As with attitudes and preferences, we believe that complexity expectations are often constructed rather than retrieved, using both internal and external inputs—namely, (1) consumers' relevant prior knowledge, (2) the observable experiences of others, and (3) communications from marketers (Fournier and Mick 1999; Oliver and Winer 1987; Shih and Venkatesh 2004). Although complexity expectations may arise from several different factors, we focus on two relevant factors: product usage information provided by the retailer (a marketer action) and consumers' prior expertise (a latent consumer characteristic). According to Oliver and Winer (1987), both of these types of information are hypothesized to influence the formation of expectations. Notably, these two factors may also interact so that an expert consumer may be differentially affected by marketer/retailer-provided information compared with a novice consumer.

Complexity expectations may arise from marketer communications. We have discussed various conceptual ways that expectations can affect product evaluations. In many practical ways, managers can influence complexity expectations, such as when they show a product being used in commercials, describe usage instructions in online formats, or use other high-involvement sales channels (e.g., sales representatives, infomercials). A tactic that marketers increasingly employ is the in-store product demonstration (Raphel and Raphel 1994). Product demonstrations reduce the consumer's perceived uncertainty before purchase and have become an important marketing-mix variable in the automobile and computer industries (Heiman and Muller 1996) and in business-to-business domains (Gopalakrishna et al. 1995). Product demonstrations can be both benefit oriented (e.g., let consumers gain information about experience attributes) and process oriented (e.g., teach consumers how to set up or use a product feature). Heiman and Muller (1996) suggest that the benefit-oriented value of demonstrations should be either positive or negative depending on situational characteristics (e.g., the type of product [or product attribute] and its inherent risk). Although Heiman and Muller do not address postpurchase evaluation and satisfaction, they show strong empirical evidence that the efficacy of demonstrations on purchase probabilities (they focus on benefit-oriented demonstrations) varies with consumers' product knowledge. Thus, it is easy to argue that the process-oriented value of demonstrations similarly depends on situational characteristics; in this case, the value of the demonstration depends on whether consumers set high or low expectations for their own difficulty in learning about and using the product.[3] In line with Heiman and Muller's findings, a consumer's expertise may be an important factor in the expectations he or she sets both before and after the demonstration.

Thus, we posit that consumers' product category expertise affects both initial complexity expectations and the extent to which these expectations are influenced by product demonstrations. Research in expertise suggests the likely pattern of this influence. Specifically, novices may not be able to imagine what they do not know (Kruger and Dunning 1999) and thus may be more positive when there is no demonstration. When a demonstration is provided for an unfamiliar product, the product demonstration may increase both the specificity and the negativity of the novices' expectations (McGill and Iacobucci 1992). Effectively, the demonstration may make the technological intricacies required to master the new product highly salient, increasing the novices' expectations (perhaps unrealistically) for difficulty of use. Conversely, expert consumers often approach new product use confidently (Alba and Hutchinson 2000). Because they rely heavily on their own prior experiences, experts often pay little attention to any usage information the firm provides (Wood and Lynch 2002). Formally, we predict the following:

---

[2] It can also be argued that benefit-oriented expectations become more complex with new product categories, and though this topic is still an important area for further research, it is addressed in the work of Meyers-Levy and Tybout (1989), Moreau, Lehmann, and Markman (2001), and Hoeffler and Ariely (1999).

[3] We thank an anonymous reviewer for articulating this point.

> H_1: Product demonstration (a) increases novice consumers' expectations that the product will be difficult to use but (b) does not affect expert consumers' expectations.

We predict that demonstrations will not influence experts' expectations, because experts base their expectations largely on information stored in long-term memory, the ambiguity of which has been discounted (Oliver and Winer 1987). Such discounting diminishes the biasness of the expectation, and thus disconfirmation is less likely to occur. Conversely, novices are more likely to hold biased expectations, which can be disconfirmed in one of two ways: (1) Those exposed to a demonstration could experience positive disconfirmation ("better than expected") at first use, and (2) those not exposed to a demonstration could experience negative disconfirmation ("worse than expected"). Formally,

> H_2: (a) Novices who experience a product demonstration report more positive disconfirmation of complexity expectations (better-than-expected ease of use) after first use of the innovation than novices who do not experience a demonstration, whereas (b) product demonstration does not affect expert consumers' disconfirmation of complexity expectations.

Step 3: The Influence of Disconfirmation on Experienced Emotion

However, it would be hasty (and possibly ill-advised) to assume that support for H_1 and H_2 suggests that in-store product demonstrations are a bad idea, having no influence on experts and a negative influence on novices. Because of the relationship between disconfirmation of expectations and emotion, the potential negative influence of product demonstrations on novices may ultimately lead to more positive results. By depressing expectations of an easy experience, novices who receive a demonstration may subsequently be positively surprised by the actual usage experience, seeing it as "not so bad after all," a situation that can create positive emotions. The E^3 model suggests that consumers who experience different disconfirmation of their usage expectations have different emotional responses.

Why do consumers experience emotions in using complex products? When people are confronted with a learning task, they develop goals and then monitor progress toward those goals. Emotions are mechanisms that communicate important information relative to expected progress toward goal achievement (Bagozzi, Gopinath, and Nyer 1999; Luce, Bettman, and Payne 2001); as such, they are likely to occur spontaneously when consumers first use a new, complex product. Specifically, negative emotions occur when expected progress toward the activated goal is impeded, and positive emotions occur when expected progress toward the activated goal is accelerated or when the goal is attained (Carver and Scheier 1990; Luce, Bettman, and Payne 2001).

Although consumption emotions can be driven by actual product performance (Oliver 1993; Phillips and Baumgartner 2002), research also suggests that consumption emotions are a function of disconfirmation of expectations (Westbrook and Oliver 1991). In the introduction of relatively noninnovative new products (e.g., batteries, hotels [Rust et al. 1999; Voss, Parasuraman, and Grewal 1998]) or in domains in which consumers have experience (e.g., current subscribers of an interactive television entertainment service [Bolton and Lemon 1999]), consumers may be able to generate accurate usage expectations, and disconfirmation is not likely to occur. Although uncertainty remains about the quality of the consumption experiences in these cases, the consumer has little uncertainty about the difficulty in consuming the product. These situations can be contrasted with those in which consumers are actively learning the new category (e.g., working with new self-service technologies [Meuter et al 2005], participating in coproduction [Bendapudi and Leone 2003]) in which real gaps are likely to exist between consumers' expectations of the ease/difficulty of goal attainment and their actual experience. Appraisal of progress, especially given unexpected difficulty or ease, typically generates an emotional response (Smith and Ellsworth 1985). Thus, we propose the following:

> H_3: Positive disconfirmation of complexity expectations is correlated with positive emotion during use, and negative disconfirmation is associated with negative emotion during use.

Step 4: Influence of Emotion on the Evaluation of Innovative Products

Can emotions affect product evaluations? It could be reasonably argued that emotions that arise from "temporary" learning or trial periods are unlikely to influence more long-lasting product evaluations. However, recent research in emotion shows that affective influence is often stronger and more far-reaching than previously considered (e.g., Bagozzi, Gopinath, and Nyer 1999). Several recent articles have shown the significant link between consumption emotions and satisfaction (Oliver 1993; Phillips and Baumgartner 2002); most empirical findings indicate that both positive and negative emotions influence satisfaction in their respective directions (Schwarz and Clore 1983). Yet demonstrations of emotional influence on product evaluation have been mixed. In a longitudinal study, Dubé and Morgan (1996) demonstrate that retrospective measures of emotion do not influence product—or, in their case, service—evaluation.

Although we posit that the addition of emotional responses benefits traditional models of diffusion, the opposite could also be suggested. A reasonable argument may be that emotions result from the adoption process but that their effect on evaluations is negligible compared with traditional models that capture both net benefits and product complexity. As we noted, net-benefit models have a strong track record of effectively predicting innovation evaluations, sometimes accounting for more than 50% of explained evaluative variance (e.g., Moreau, Lehmann, and Markman 2001).

We hypothesize that because emotions result from disconfirmation of usage expectations, they act as important signals for the consumer regarding his or her own experience. This hypothesis is akin to other examples of "affect as information" in which evaluations of products are influenced by situational mood states (Pham 1998). This emotional influence may operate in addition to other sources of emotion, such as the utility derived from experienced benefits. Because usage expectations are different from consumers' expectations about the product's ultimate benefits, focusing instead on early usage experience (e.g., the rate at which they can achieve those benefits, the ease of learning, and their own affective response), the emotions derived from usage expectations may be distinct from emotions experienced from gaining product benefits.

The changing influence of emotions over time. It is important to consider the potentially dynamic influence of emotion at distinct periods in the adoption process. Whereas the dominant satisfaction paradigm assumes that consumers use static standards to make satisfaction judgments, Fournier and Mick (1999) find strong evidence that standards are often malleable. In the early stages of product learning, we also expect to observe dynamic changes in usage difficulty expectations as consumers gain direct experience with the new product and update their priors accordingly. In particular, inexperienced consumers, who may have been surprised by the difficulty or ease of their initial usage experiences, are likely to use their newly acquired direct experience to adjust their expectations for their next use. Although the adjustment may not be accurate (e.g., these consumers may either over- or underadjust), basic tenets of learning hold that over time and experience, usage expectations should stabilize. Over time, we expect emotional influence to be the greatest during the early learning stages of product usage, when the disconfirmation gap between expectations and actual experience is at its largest; then, as learning occurs, we expect this gap to shrink, thus decreasing the intensity of the emotions and their correlation with evaluations. Formally,

> H_4: Both positive and negative emotions (a) are significant predictors of evaluation at the initial introduction of the new product (before first use), (b) are significant predictors of evaluation at the first use of the new product, and (c) are not significant predictors of evaluation after repeated use of the product.

Shih and Venkatesh (2004) call specifically for experimental and longitudinal work to address usage diffusion elements. In Study 1, we use a controlled laboratory setting to test our hypotheses. In Study 2, we use a quasi experiment over a three-month period to explore these hypotheses further in a real-world setting.

Study 1

Methods

Choice of innovation stimuli. Three requirements guided our selection of a new product: (1) It had not been used by the participant population, (2) there was natural variation in prior experience with a previous version, and (3) learning was necessary to achieve basic performance goals. We selected a new personal digital assistant (PDA), the Palm Zire—specifically, the use of the handwriting-recognition software. The Palm Zire offers both a calendar function and an address function, and to use each of these functions, users must learn the proprietary Graffiti handwriting-recognition software.[4] Learning to use the Graffiti software was chosen as a task that was both complex and fundamental to the operation of the new product and that could be viably learned and tested within the parameters of the laboratory setting.

Design and procedure. The study was a 2 (product demonstration: no versus yes) × 2 (prior PDA experience: no versus yes) between-subjects design. Participants were 175 undergraduate students at a large midwestern university who participated in the hour-long study for course credit. After a brief overview of PDAs and the two required tasks, each participant was given the experimental materials that detailed how to use the Graffiti characters. Next, participants in the randomly assigned "demonstration" condition received the demonstration provided by the same experimenter using the same script.[5]

All participants then read additional details about the Palm Zire and answered the preusage dependent measures. Next, the experimenter delivered a Zire to each workstation. Participants were given step-by-step instructions and asked to complete the first task of entering an appointment. Participants showed completed work to the experimenter and answered the next set of dependent measures. After this, the participants proceeded to the second task, namely, entering an address. Participants then answered a final set of dependent measures; before leaving, they were thanked and debriefed.

Independent and Dependent Measures

Prior category experience. Participants indicated their experience level with PDAs on a seven-point scale. Approximately half indicated no experience or a very low level (i.e., a 1); thus, we used a median split to classify participants as inexperienced versus experienced.

---

[4] There is an alternate way to enter information using a simulated keyboard, but the experimental instructions required participants to use the Graffiti software.

[5] In the demonstration, the experimenter held up the Zire, showed the participants where the stylus was located, and indicated how to tap the main menu buttons to find the appropriate tasks. The experimenter then put a transparency of the sheet detailing the Graffiti characters on an overhead projector and briefly demonstrated how and where to draw the letters on the screen of the Zire.

Product-related expectations. With the exception of questions on prior PDA usage and demographics, we employed the same set of dependent measures at each of the three points of interest during the experiment. To allow participants to communicate any type of product-related expectation, we employed an open-ended written protocol both before and after product trial ("As you anticipate using the Palm Zire PDA to schedule an appointment, please describe the expectations that you have about your first experience using the Palm Zire PDA"). An examination of the responses indicated that almost all contained information on the anticipation of and reactions to the Zire's ease/difficulty of use. Two coders, who were blind to the condition of the participants, coded the written responses on two seven-point scales that captured expectations of difficulty in learning and using. After each subsequent use, the coder rated the written protocols on two additional scales that captured how the experience compared with expectations (much more difficult to learn, much more difficult to use). The coders' reliability was high (correlation on all scales >.74); thus, we averaged their responses. Furthermore, the correlation between each pair of items was also high (each pair >.88); thus, we averaged each pair of items to form a measure of usage expectations for Time 1 and a measure of disconfirmation of usage expectations for Times 2 and 3. We combined the coded measures for Time 1 with the traditional complexity measures that Rogers (1996) used (we describe this further in the next section) to form an overall expectation score.

Complexity expectations. Following Rogers's (1996) definition of complexity, participants indicated on three seven-point scales how difficult they expected/perceived the Zire to be to use, how long it would take to learn to use, and how much of a challenge it would be to use (at each period, $\alpha > .90$); higher numbers indicated greater expected difficulty. At Time 1 (preusage), we combined these three self-report measures with the two items derived from the written protocols to create an overall index of expected usage difficulty ($\alpha = .92$).

Disconfirmation of complexity expectations. A standard method for assessing disconfirmation is to measure participants' retrospective assessments of whether the product or service was better or worse than expected (e.g., Oliver 1993; Phillips and Baumgartner 2002). Effectively, our coding of the written protocols at Times 2 and 3 captures a similar appraisal of this comparison in a more open-ended manner. In addition, we subtracted usage expectations at time t from usage expectations at time t + 1. We combined this difference with the two coded items.

Emotions. To assess participants' emotions at each stage in the adoption process, we employed a modified version of Izard's (1977) differential emotions scale (DES-II); we balanced the scale's positive and negative items as Richins (1997) suggests. Similar to previous consumption emotion research (e.g., Phillips and Baumgartner 2002), we found a two-factor solution at each point in time using the modified scale, with the two factors representing the positive and negative emotions. We combined the seven positive and seven negative items to create a positive and a negative emotion variable for each point in time (see Figure 2).

Product evaluations. We used a five-item scale with seven-point items to assess participants' evaluations of the Zire at each of the three points in time ("As an overall evaluation of this particular PDA, I think I will 'dislike it very much/like it very much'"; "I think this particular PDA and its features are good": "strongly disagree/strongly agree"; "I am very pleased that I will have to try this particular PDA in this study": "strongly disagree/strongly agree"; "I would like to use this PDA on a regular basis": "strongly disagree/strongly agree"; "Would you recommend this PDA to your friends?": "definitely would not/definitely would"). At each point in time, all five items loaded on a single factor, and thus we summed the five items to form a single evaluation measure for each point in time (all three $\alpha > .89$).

Covariates. We included compatibility and net benefits as covariates that have been shown to influence new product success. In line with the work of Rogers (1996), we

FIGURE 2
Emotion Response Scale (Modified from Izard's DES II Checklist)

The DES Emotions Checklist. Record the emotions you are experiencing now by circling the response that best fits how you truly feel.

Not at all happy	1	2	3	4	5	6	7	Very happy
Not at all frustrated	1	2	3	4	5	6	7	Very frustrated
Not at all sad	1	2	3	4	5	6	7	Very sad
Not at all joyful	1	2	3	4	5	6	7	Very joyful
Not at all angry	1	2	3	4	5	6	7	Very angry
Not at all confused	1	2	3	4	5	6	7	Very confused
Not at all motivated	1	2	3	4	5	6	7	Very motivated
Not at all determined	1	2	3	4	5	6	7	Very determined
Not at all excited	1	2	3	4	5	6	7	Very excited
Not at all scared	1	2	3	4	5	6	7	Very scared
Not at all bored	1	2	3	4	5	6	7	Very bored
Not at all irritated	1	2	3	4	5	6	7	Very irritated
Not accomplished	1	2	3	4	5	6	7	Very accomplished
Not at all proud	1	2	3	4	5	6	7	Very proud

measured compatibility with three summed scale items to indicate how significantly participants would need to change the way they currently schedule appointments, keep track of addresses, and manage their information (at each period, $\alpha > .74$). Two scale items measured net benefits of the Zire (the perceived functional benefits less associated risks), and two other scale items measured participants' evaluations of the Zire's appointment and address functions. The four items loaded on the same factor ($\alpha > .78$), and we summed them at each of the three points in time.

Results

The factors influencing initial complexity expectations. We used a two-way analysis of variance (ANOVA), with experience and demonstration as the independent factors, to test H_1. The results indicate both a main effect of experience ($F(1, 172) = 19.62$, $p < .0001$) and the predicted interaction between the two factors ($F(1, 172) = 5.21$, $p < .02$). In support of H_{1a}, inexperienced participants who received the demonstration had significantly higher expectations of difficulty than those who did not ($F(1, 172) = 3.96$, $p < .05$; contrast: $M_{demo} = 20.9$ versus $M_{no\ demo} = 18.5$). However, the demonstration did not significantly influence those with prior experience ($F(1, 172) = 1.83$, $p > .10$; $M_{demo} = 14.8$ versus $M_{no\ demo} = 16.6$), in support of H_{1b}.

The disconfirmation of complexity expectations at first use. The negative expectations that the product demonstration created for inexperienced consumers may not be ultimately detrimental to product evaluations. With the measure of disconfirmation as the dependent variable, we used a two-way ANOVA to test H_2, which revealed the predicted interaction ($F(1, 172) = 14.23$, $p < .001$; see Figure 3, Panel A). Inexperienced participants who received the demonstration experienced the least negative disconfirmation of expectations (i.e., the most positive disconfirmation, finding the Zire easier to use than expected) compared with their counterparts who did not receive the demonstration ($F(1, 172) = 22.53$, $p < .0001$; contrast: $M_{no\ demo} = 4.5$ versus $M_{demo} = -2.6$, $p < .0001$), in support of H_{2a}. Again, the demonstration did not significantly influence the degree of disconfirmation for those with prior experience ($F(1, 172) = 1.09$, $p > .10$; $M_{demo} = 4.2$ versus $M_{no\ demo} = 2.3$), in support of H_{2b}.

Impact of disconfirmation on experienced emotion and evaluation at first use. If disconfirmation influences emotion (and subsequently product evaluation), we would expect to observe an experience × demonstration interaction for emotion that mirrors the pattern tested in H_2. The predicted interaction is significant ($F(1, 170) = 4.90$, $p = .02$; Figure 3, Panel B). Experienced participants' emotions were not influenced by the product demonstration, yet inexperienced participants were strongly influenced by the demonstration; those who had a demonstration reported more positive emotion ($M_{demo} = 32$) than those who did not have a demonstration ($M_{no\ demo} = 27$).

Another way to test H_3 is through regression analysis by examining the comparative influence of complexity expectations on positive and negative emotion at all three periods (before use, at first use, and after repeated use). At each point in time, we ran two regressions with positive and negative emotions as the dependent variables. Predictors included the two independent factors (experience and demonstration), their interaction, the measure of usage expectations, and the covariates (compatibility and net benefits). At Times 2 and 3, we also included as predictors the disconfirmation of usage expectations. Perceptions of the Zire's usage difficulty and its net benefits significantly predicted emotions at each point in time. As Table 1 shows, complexity is a significant driver of emotion at Time 1 ($B_{positive} = -.31$, $p < .001$; $B_{negative} = .20$, $p < .05$). Before use, net benefits influence only positive emotion ($B = .18$, $p < .01$), not negative emotion. Importantly, this pattern also holds at Time 2 and Time 3, augmenting the support for H_3.

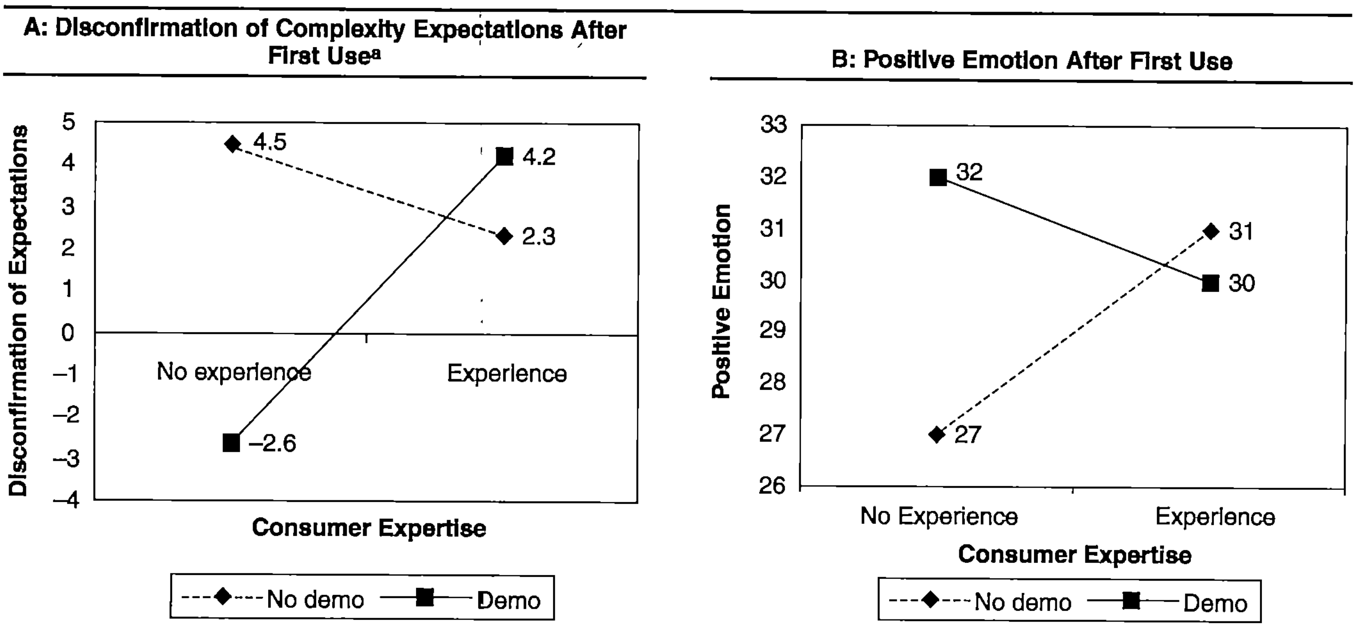
Similar to H_3, we test H_{4a-c} with regressions at each period. For each period, we ran a regression using evaluations of the Zire as the dependent measure. Predictors included the two independent factors, their interaction, the measures of usage expectations, the covariates (compatibility and net benefits), and positive and negative emotions. At Times 2 and 3, we also included as predictors the disconfirmation of complexity expectations. Before first use, positive emotion was a significant driver of evaluation ($B = .41$, $p < .001$), as were net benefits ($B = .27$, $p < .001$; see Table 2). However, negative emotions were not significant predictors of evaluation ($B = -.10$, $p > .05$). This result is notable, and we discuss it further in Study 2. After first use, both positive ($B = .33$, $p < .001$) and negative ($B = -.22$, $p < .01$) emotions were significant drivers of product evaluations, as were net benefits ($B = .25$, $p < .001$). Finally, after repeated usage, positive ($B = .47$, $p < .001$) and negative ($B = -.24$, $p < .001$) emotions were still both significant drivers of product evaluation, as were net benefits ($B = .25$, $p < .001$). Thus, the influence of emotion can be observed at all points in time, in support of H_{4a} and H_{4b} but not H_{4c}.

Discussion

The results from Study 1 demonstrate that a consumer's expectations about the difficulty he or she will experience in using a new product can be influenced by both marketer-directed communications and by his or her own prior experience. These usage expectations and their (dis)confirmation were important factors that predicted emotions and drove their changes over the course of the usage experience. In turn, emotions predicted product evaluations at all points in time, even after we controlled for the other diffusion variables.

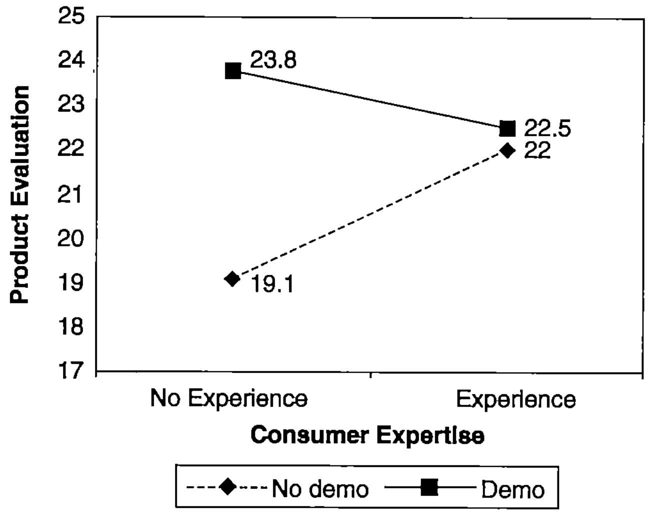
Note that only positive emotions predicted evaluations toward the Zire before its use, whereas both positive and negative emotions influenced evaluations after use. Before use, any threats to expected goal progress were only hypothetical; perhaps the negative emotions generated before use of the Zire were not substantial enough to influence evaluations. Only when the expected progress was actually impeded did negative emotions influence evaluation. This phenomenon might be conceptualized as a "suspension of judgment," in which the consumer holds doubts or negative feelings at bay in the initial assessment of the product. In other words, the consumer may experience negative emo-

FIGURE 3
Expectation Disconfirmation, Experienced Emotions, and Product Evaluation After First Use



A: Disconfirmation of Complexity Expectations After First Use[a]

B: Positive Emotion After First Use



C: Product Evaluation After First Use

[a] Larger numbers indicate greater negative disconfirmation (worse-than-expected usage complexity).

tions when considering the learning/usage difficulties but may not incorporate these initial emotions in evaluating the product. This is a notable and somewhat counterintuitive finding, given that consumers pay disproportionately more attention to negative attribute performance in postuse product evaluations (Mittal, Ross, and Baldasare 1998).

Innovation Continuity and Complexity Expectations

In Study 1, the complexity expectations of the inexperienced consumers were the most influenced by the product demonstration. As such, these consumers were subject to the greatest amount of disconfirmation and, consequently, produced greater emotional reactions to the Zire. In this population, inexperienced participants had no experience with the PDA category. Discontinuous innovations have been defined as those that do not fit neatly into an existing category (Lehmann 1994), and for consumers with no prior PDA experience, there is likely to be no substantial PDA category representation existing in memory. For consumers with PDA experience, however, such a representation exists. As such, the Zire could be considered a discontinuous innovation for those without PDA experience but a con-

TABLE 1
Study 1: Predictors of Emotions (Standardized Estimates)

	Emotions					
	Before Use		After First Use		After Second Use	
Predictors	**Positive**	**Negative**	**Positive**	**Negative**	**Positive**	**Negative**
Prior PDA experience	.03	–.12	.08	–.10	.07	–.02
Product demonstration	.10	–.04	.01	.01	.18*	–.07
Experience × demo	–.07	.01	–.04	–.04	–.12	.06
Product-Related Expectations						
Net benefits	.18**	.10	.37***	–.08	.16*	–.04
Compatibility	–.08	–.11	–.11	–.04	–.01	–.04
Usage (difficulty)	–.31***	.20*	–.52***	.88***	.42***	.68***
Disconfirmation of usage expectations			.04	.19*	–.17*	.07
R^2	.10	.12	.33	.59	.31	.54

*p < 05.
**p < .01.
***p < .001.

TABLE 2
Studies 1 and 2: Predictors of Evaluations of the Innovation (Standardized Estimates)

	Evaluations					
	Before Use		After First Use		After Second Use	
Predictors	**Study 1**	**Study 2**	**Study 1**	**Study 2**	**Study 1**	**Study 2**
Prior experience	–.06	.13	.01	.13	–.07	–.01
Product demonstration	.04	.01	–.01	.21	–.01	.01
Experience × demo	–.04	–.04	.03	–.28	–.01	–.04
Product-Related Expectations						
Net benefits	.27***	.36***	.25***	.13	.25***	.38***
Compatibility	.08	.02	.04	.07	–.01	.18
Usage (difficulty)	–.16*	–.18*	–.23*	–.51***	–.26***	–.21*
Disconfirmation of usage expectations			–.16*	.36*	.05	–.03
Consumption Emotions						
Positive	.41***	.35***	.33***	.35***	.47***	.22*
Negative	–.10	–.08	–.22**	– 24*	–.24***	–.11
R^2	.33	.52	.65	.51	.71	.46

*p < .05.
**p < .01
***p < .001.

tinuous innovation for those with it. Because the literature has documented the effect of continuity on responses to new products (e.g., Moreau, Lehmann, and Markman 2001; Rogers 1996), we designed Study 2 to control for the continuity of the innovation, while assessing differences in product experience.

With a discontinuous innovation, the magnitude of the gap between expected and actual goal progress is likely to be significant because the consumer has no direct category experience on which to base expectations (Oliver and Winer 1987). As a result, the emotions that ensue are likely to be more powerful and predictive of overall evaluations. However, with a continuous innovation, the size of the gap is likely to be smaller. In these situations, all consumers are likely to have some direct experience with existing products in the category, but the level may vary across consumers. Although their existing knowledge may not be perfectly calibrated, it is likely to provide a more accurate prediction of the actual usage experience than if it were unavailable. In Study 2, we test this boundary condition to determine whether the E^3 model (H_1–H_4) holds for a continuous innovation.

Study 1 also involved the typical level of artificiality that is characteristic of lab studies, and its controlled nature dictated an accelerated usage schedule. As such, participants had no discretion over the timing of the use of the product. In both academic (e.g., Rogers 1996) and practitioner-oriented (e.g., *The Economist* 2004) literature, there is evidence that the expected usage difficulty of a product causes consumers to delay a purchase. Thus, we might expect differential complexity expectations to affect usage timing even in situations in which the consumer does not make a purchase decision, such as with a forced adoption. Formally,

> H_5: Expectations of complexity are positively correlated with the delay in initial usage.

In Study 2, we use a quasi-experimental methodology over an extended period to test H_1–H_5 in the context of a continuous innovation. Whereas we predict that complexity expectations are less influential for a continuous innovation, we expect (and demonstrate) similarities between Study 1 and Study 2 in the factors that influence overall evaluations. As we observe, the pattern of emotions predicting evaluations of the discontinuous innovation at all three times is almost identical to the pattern predicting evaluations of the continuous innovation.

Study 2

Method

As in Study 1, several criteria guided our selection of stimuli: (1) The product had to be a continuous innovation for all participants, with natural variation in usage experience among the set, and (2) the product had to require some learning for all participants to achieve its basic performance goals. We selected PageOut, a new course management software program from McGraw-Hill. A pretest confirmed that there was variance in participants' experience in using online course management tools but that all potential (student) participants were highly familiar with different types of software, with Web page interfaces, and with computers. As a cover story, students were told that they had been volunteered to beta-test the software and would be asked to assess PageOut three times throughout the semester.

Design and Protocol

Undergraduates who were enrolled in three sections of an introductory marketing course participated in this 2 (high versus low experience) × 2 (demonstration versus no demonstration) between-subjects study throughout the semester as part of their curriculum. Although 123 students were enrolled in the three sections, 17 failed to complete all aspects of the study and were subsequently deleted from further analyses. The students were not randomly assigned to a section, but the course prerequisites, the professor, the class format, and the course content were held constant across all three sections.

On the first day of class, the professor introduced PageOut to the students in all three sections using a consistent script and provided all students with the same handout containing step-by-step instructions. In one randomly selected section, the professor used the computer podium to log on to the Web site and to demonstrate how to create a student profile. The other two sections received just the verbal introduction and the instructional handout for this task. This operationalization of the demonstration differed from Study 1 in that it used a different medium (computer), used a different instructional material for the Web site's use, and was conducted in a group rather than individually.

Before use (Time 1). On the first day of class, immediately following the introduction, students reported the emotions they experienced while the professor was introducing PageOut. Students then completed demographic and experience measures, followed by the dependent measures. Finally, students were given an assignment to log on to PageOut and register by the next class. A small extra credit incentive was provided. The administrators' software enabled the professor to record students' log-on and registration times.

After the first use (Time 2). On the second day of class, students reported the emotions they felt while logging on and registering with PageOut and responded to the dependent measures.

After extended usage (Time 3). Over the course of the semester, students used PageOut to take online quizzes, check exam grades, download class notes, and communicate with the professor. At a mandatory class meeting shortly before the end of the semester, students reported their current emotions and completed the final set of dependent measures.

Dependent Variables

Expectations, emotions, and evaluations. Measures of product-related expectations, emotions, and evaluations were similar to those used in Study 1, with one exception. In an effort to minimize the time required to collect the data, no open-ended protocols for usage expectations were employed.

Log-on time. As we noted previously, the administrators' software enabled the professor to track the time that elapsed from the time each student left class to the time that he or she registered on the Web site.

Prior experience. Experience with online course management tools was measured using two items. A median split identified students as experienced or inexperienced.

Results

Initial usage expectations and timing of initial use. H_5 predicted that higher expectations of usage difficulty would be associated with delays in product usage. The simple correlation between participants' perceptions of usage difficulty and the hours they allowed to elapse before their initial registration suggests that such a mechanism was at work ($r = .25, p < .01$).

The influence of experience and demonstration on initial complexity expectations. We tested H_1 using a two-way ANOVA. As in Study 1, there was a significant main effect of experience ($F(1, 102) = 7.30, p < .01$). The pattern of data was consistent with the first study; participants with more experience with course management software expected their initial use of the PageOut software to be less difficult than did those with less experience ($M_{experienced}$ = 4.9 versus $M_{inexperienced}$ = 6.8).

The accuracy of complexity expectations for a continuous innovation. After participants logged on and registered, how did their initial usage expectations compare with their actual experience? Given the continuous nature of the innovation, we expected that H_2 would be attenuated in this study. Accordingly, in contrast to the findings in Study 1, there were no significant effects of prior experience or demonstration on the degree of disconfirmation of complexity expectations. The means for all four groups were close to zero, indicating that participants were relatively well calibrated in their expectations.

Consumption emotions and product evaluations. As we expected, given participants' greater accuracy in complexity expectations, disconfirmation has a much smaller influence on experienced emotion at each of the three periods. We replicated the experience × demonstration interaction for felt emotion at first use observed in Study 1 (using standardized measures, $B = -.21$, $p < .05$, $r = .19$). Less-experienced participants experienced more positive emotion at first use when they received a demonstration ($M = .31$) than when they did not receive one ($M = -.24$), whereas more-experienced participants were not affected similarly by the demonstration ($M_{demo} = -.12$ versus $M_{no\ demo} = .13$). However, as we expected with the small degree of disconfirmation, the experience × demonstration interaction for Time 2 evaluations of the innovation was not significant ($F = .67, p > .2$).

There are dramatic similarities between the Study 1 and the Study 2 findings in the factors that influence overall evaluations. The pattern of results that predict evaluations of the Zire across time is almost identical to that of PageOut (Table 2). In both studies, only positive emotions predict evaluations before use, which might be interpreted as an initial suspension of judgment during the product introduction such that negative emotions, though they are experienced, are held at bay. However, both positive and negative emotions significantly influenced evaluations after use in both studies, in support of H_{4b}. Finally, H_{4c} predicted that the influence of emotions would wane over time. Unlike Study 1, in which the impact of emotions remained strong during the duration of the one-hour session, the impact of emotions declined over the semester for PageOut, in support of H_{4c}. Taken together, these findings suggest that the diminishing influence of emotions can be observed only over longer time horizons.

Study 2 demonstrates that the pattern of results from Study 1 holds in a real-world quasi experiment but that emotional influence is not as extreme when the product is less innovative. However, even when emotional influence was smaller as a result of more accurate complexity expectations, the inclusion of emotion still increased the predictive power of the traditional diffusion model.[6] Finally, this study shows that complexity slows consumers' time to trial (H_5).

General Discussion

Theoretical Contribution

The E^3 model. By demonstrating how affective responses during new product learning influence innovation evaluations over time, we extend previous work in both satisfaction and diffusion. The E^3 model provides a parsimonious description of how emotions are created in the early use of complex products and how they may influence product evaluations beyond diffusion's traditional focus on net benefits. These results emphasize the dynamic nature of both emotion and evaluation in innovation diffusion.

Adding emotion to innovation frameworks. Refining models of innovation is easily justified by the importance of new product success to company profitability. Prior research on new product adoption focuses largely on cognitive processes, despite the increasing knowledge of the role of emotion in product consumption. We demonstrate that emotions resulting from early experience with the product (which therefore might be dismissed as ephemeral) can have a lasting influence on diffusion.[7]

Exploring and expanding recent innovation research. The use-diffusion model (Shih and Venkatesh 2004) describes four patterns of consumer use of new products based on the rate and variety of consumer use. Shih and Venkatesh (2004) theorize that usage diffusion depends on (1) characteristics of the individual or household (e.g., innovativeness, resources, prior experience), (2) characteristics of the innovation (e.g., complexity, technological sophistication, existence of complementary products), and (3) characteristics of the environment (e.g., communication intensity, technological access, media exposure). The current research expands the knowledge of how these factors may interact by examining how the characteristics of the environment (how the innovation is promoted to the consumer), the product (complexity), and the consumer (prior expertise) combine to create a successful or unsuccessful initial product experience.

In addition, the importance of the phenomena we report herein is reflected in recent work on feature fatigue. Thompson, Hamilton, and Rust (2005) show that consumers assign more weight to product capabilities (additional features) in their evaluations before use than after use and less weight to product usability before use than after use. Thus, they show that consumers are often miscalibrated in their expectations about product complexity and usability before use. Our research shows that inaccurate and overconfident expectations about usability (e.g., when consumers seek excessive features in prospect) may create situations in which negative emotions have a powerful influence on product evaluations and, thus, in retrospect, satisfaction.

---

[6] To assess this contribution, we compared the explanatory power of the model shown in Table 3, Panel B, with that of a reduced model that does not include the two emotion variables ($R^2 = .38$). The resulting F statistic ($F(2, 95) = 9.13$, $p < .01$) demonstrates that the addition of emotions to the standard model of preference makes a significant contribution.

[7] Researchers have long advocated the addition of affective elements to primarily utilitarian-based models of product evaluation (Derbaix and Pham 1991). However, many researchers view emotion as another viable end-state benefit similar to utilitarian benefits. For example, Havlena and Holbrook (1986) note that consumers might buy a luxury car for the pleasure of feeling superior. We do not investigate emotions as independent benefits in this work but rather as a moderator of product learning and acceptance. Although emotions can color utilitarian attribute evaluation (Smith and Ellsworth 1985), they can also have a more global influence on product experiences (Cohen and Areni 1991).

Managerial Implications

Our results show that the predictive power of traditional models of innovation adoption can be improved by the addition of emotional responses. Importantly, a substantial proportion of consumers' emotional responses appear to be a function of disconfirmed complexity expectations, not simply a result of product benefits. This tells marketers two things. First, emotional influence can be assessed by measuring disconfirmation of usage expectations, and second, emotional responses can be influenced by changing consumers' complexity expectations before use. As we noted previously, marketers can influence complexity expectations when they show product use in commercials, describe usage instructions in online formats, or use other high-involvement sales channels (e.g., sales reps). We tested product demonstration as one way to influence expectations. Product demonstrations are believed to be beneficial to both marketers and consumers but expensive to execute, and marketers do not have a clear idea of how they truly affect consumers (Ailloni-Charas 2000). Our research suggests that training frontline sales personnel is important; the salesperson may be tempted to insist that a technological product is easy to learn to close the sale, yet the mismanagement of usage expectations can lead to high rates of product return in the long run. Promising "easy" use may promote trial but sabotage product evaluation and diffusion.

Thus, marketers must seek a balance in which consumers are encouraged to try new products but not through promises of ease that create unrealistic expectations for early use. For example, we can compare the recent strategic actions of Royal Philips and Best Buy to observe two opposite outcomes of influencing consumers' complexity expectations. Best Buy's Geek Squad (currently in 642 stores) is a team of professionals, attired in stereotypically nerdy uniforms (black polyester pants, white shirts, and clip-on ties), who make house visits to help consumers learn complicated product tasks, such as how to program an iPod or connect a TiVo system (Deam 2004). Rather than downplaying the complexity of products sold in the store, Best Buy leverages these perceptions by offering (for a fee) a way to shift the learning costs to a team of prototypically smart people. Not only does this generate revenue, but our research would predict that the Geek Squad also sends a signal, albeit humorous, to consumers that product learning is not always easy; with such a signal, consumers may subsequently experience little negative disconfirmation in early usage and be positively surprised if they successfully achieve product use on their own. Conversely, Philips's strategy of promoting itself as a provider of user-friendly gadgets (*The Economist* 2004) may be less optimal for the company. By launching a $100 million ad campaign to promote its technology as "easy to use" (Endt 2004), Philips may lower consumers' preuse complexity expectations and subsequently strengthen the negative emotions that occur in the early learning phase of use. Most learning requires some setbacks, and reducing consumers' expectations of setbacks may have a negative impact on long-term product and brand evaluations.

Limitations and Further Research

A limitation of the current research is our use of forced adoption in the two studies. Participants in both studies did not choose to adopt the new products, though they had an outlet to report their opinions and evaluations at multiple points in the usage period. Although this situation was chosen for practical reasons, on reflection, we question whether forced adoption has a positive side in offering a better and more conservative test of the influence of complexity expectations and emotion. When consumers choose to try/buy a product innovation, there are certain emotional and evaluative processes that arise solely from the evaluation of the choice, such as cognitive dissonance. Dissonance may amplify emotions that arise from disconfirmation of usage expectations. Thus, emotional responses may be heightened in voluntary adoption, creating even stronger influence of emotions on innovation evaluations.

There are many other diverse avenues for further research suggested by our results. First, the E^3 model advocates the inclusion of both complexity expectations and emotions (positive and negative) in existing frameworks of evaluation, satisfaction, and diffusion. Second, these results, combined with work on regret (e.g., Tsiros and Mittal 200), begin to build a more complete picture of negative emotion after a choice is made. Luce, Bettman, and Payne (2001) show the importance of negative emotion as consumers contemplate choice. The act of making a choice is not the end of emotional influence. Consumer expectations for the usage or learning process may have a surprisingly strong impact, influencing regret over product choice dissatisfaction not only with product benefits but also with the costs to achieve those benefits in a timely manner. Can complex products make consumers feel smart because of unexpected hurdles overcome or stupid because of unexpected slowness in achievement? More important, can such emotional reactions influence the evaluation of the innovation beyond more concrete evaluative criteria? The current research suggests that they can.

REFERENCES

Ailloni-Charas, Dan (2000), "Demos, Yes, But Don't Give Away the Store," *Brandweek*, (March 13), 30.

Alba, Joseph and J. Wesley Hutchinson (2000), "Knowledge Calibration: What Consumers Know and What They Think They Know," *Journal of Consumer Research*, 27 (September), 123–56.

Bagozzi, Richard P., Mahesh Gopinath, and Prashanth U. Nyer (1999), "The Role of Emotions in Marketing," *Journal of the Academy of Marketing Science*, 27 (Spring), 184–206.

Bendapudi, Neeli and Robert Leone (2003), "Psychological Implications of Customer Participation in Co-Production," *Journal of Marketing*, 67 (January), 14–29.

Bolton, Ruth and Katherine Lemon (1999), "A Dynamic Model of Customers' Usage of Services: Usage as an Antecedent and Consequence of Satisfaction," *Journal of Marketing Research*, 36 (May), 171–87.

Cadotte, Ernest, Robert Woodruff, and Roger Jenkins (1987), "Expectations and Norms in Models of Consumer Satisfaction," *Journal of Marketing Research*, 24 (August), 305–314

Carver, Charles S. and Michael F. Scheier (1990), "Origins and Functions of Positive and Negative Affect: A Control-Process View," *Psychological Review*, 97 (January), 19–35.

Cohen, Joel B. and Charles S. Areni (1991), "Affect and Consumer Behavior," in *Handbook of Consumer Behavior*, Thomas S. Robertson and Harold H. Kassarjian, eds. Englewood Cliffs, NJ: Prentice Hall, 188–240.

Deam, Jenny (2004), "Who's Cool? It's All Geek to Gadget Users," *The Denver Post*, (November 11), D1.

Derbaix, Christian and Michel T. Pham (1991), "Affective Reactions to Consumption Situations: A Pilot Investigation," *Journal of Economic Psychology*, 12 (June), 325–55.

Dubé, Laurette and Michael S. Morgan (1996), "Trend Effects and Gender Differences in Retrospective Judgments of Consumption Emotions," *Journal of Consumer Research*, 23 (September), 156–62.

The Economist (2004), "Simplifying Philips," (June 12), 66.

Endt, Friso (2004), "It Is Rocket Science," *Newsweek*, (October 18), E8

Fournier, Susan and David Mick (1999), "Rediscovering Satisfaction," *Journal of Marketing*, 63 (October), 5–24.

Golder, Peter N. and Gerard J. Tellis (1998), "Beyond Diffusion: An Affordability Model of the Growth of New Consumer Durables," *Journal of Forecasting*, 17 (3–4), 259–80.

Gopalakrishna, Srinath, Gary L Lilien, Jerome D. Williams, and Ian K. Sequeira (1995), "Do Trade Shows Pay Off?" *Journal of Marketing*, 59 (July), 75–83.

Havlena, W.J. and Morris B. Holbrook (1986), "The Varieties of Consumption Experience: Comparing Two Typologies of Emotion in Consumer Behavior," *Journal of Consumer Research*, 13 (December), 394–404.

Heiman, Amir and Eitan Muller (1996), "Using Demonstration to Increase New Product Acceptance: Controlling Demonstration Time," *Journal of Marketing Research*, 33 (November), 422–31

Hoeffler, Steve and Dan Ariely (1999), "Constructing Stable Preferences: A Look into Dimensions of Experience and Their Impact on Preference Stability," *Journal of Consumer Psychology*, 8 (2), 113–39.

Izard, Carroll E. (1977), *Human Emotions*. New York: Plenum.

Kruger, Justin and David Dunning (1999), "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments," *Journal of Personality and Social Psychology*, 77 (6), 1121–34.

Lehmann, Donald (1994), "Characteristics of 'Really' New Products," paper presented at the Marketing Science Institute Conference, London (June 21–22).

Luce, Mary Frances, James R. Bettman, and John W. Payne (2001), "Emotional Decisions: Tradeoff Difficulty and Coping in Consumer Choice," in *Monographs of the Journal of Consumer Research, No. 1*, Deborah Roedder John, ed. Chicago: University of Chicago Press.

McGill, Ann and Dawn Iacobucci (1992), "The Role of Post-Experience Comparison Standards in the Evaluation of Unfamiliar Services," in *Advances in Consumer Research: Diversity in Consumer Behavior*, Vol. 19, John Sherry and Brian Sternthal, eds. Provo, UT: Association for Consumer Research, 570–78.

Meuter, Matthew L., Mary Jo Bitner, Amy L. Ostrom, and Stephen Brown (2005), "Choosing Among Alternative Service Delivery Modes: An Investigation of Customer Trial of Self-Service Technologies," *Journal of Marketing*, 69 (April), 61–83.

Meyers-Levy, Joan and Alice M Tybout (1989), "Schema Congruity as a Basis for Product Evaluation," *Journal of Consumer Research*, 16 (June), 39–55

Mick, David Glen and Susan Fournier (1998), "Paradoxes of Technology: Consumer Cognizance, Emotions, and Coping Strategies," *Journal of Consumer Research*, 25 (September), 123–44.

Mittal, Vikas, William T. Ross Jr., and Patrick M. Baldasare (1998), "The Asymmetric Impact of Negative and Positive Attribute-Level Performance on Overall Satisfaction and Repurchase Intentions," *Journal of Marketing*, 62 (January), 33–47.

Moreau, C. Page, Donald R. Lehmann, and Arthur B. Markman (2001), "Entrenched Knowledge Structures and Consumer Response to New Products," *Journal of Marketing Research*, 38 (February), 14–29.

Oliver, Richard L. (1993), "Cognitive, Affective, and Attribute Bases of the Satisfaction Response," *Journal of Consumer Research*, 20 (December), 418–30.

——— and John E. Swan (1989), "Equity and Disconfirmation Perceptions as Influences on Merchant and Product Satisfaction," *Journal of Consumer Research*, 16 (December), 372–84.

——— and Russell S. Winer (1987), "A Framework for the Formation and Structure of Consumer Expectations: Review & Propositions," *Journal of Economic Psychology*, 8 (4), 469–99

Olshavsky, Richard W. and Richard A. Spreng (1996), "An Exploratory Study of the Innovation Evaluation Process," *Journal of Product Innovation Management*, 13 (6), 512–29.

Pham, Michel T. (1998), "Representativeness, Relevance, and the Use of Feelings in Decision Making," *Journal of Consumer Research*, 25 (September), 144–59.

Phillips, Diane M. and Hans Baumgartner (2002), "The Role of Consumption Emotions in the Satisfaction Response," *Journal of Consumer Psychology*, 12 (3), 243–53.

Raphel, Murray and Neil Raphel (1994), "Sampling Sells," *Progressive Grocer*, 73 (January), 17.

Richins, Marsha L. (1997), "Measuring Emotions in the Consumption Experience," *Journal of Consumer Research*, 24 (September), 127–46.

Rogers, E.M. (1996), *Diffusion of Innovations*, 3d ed. New York: The Free Press.

Rust, Roland T., J. Jeffrey Inman, Jianmin Jia, and Anthony Zahorik (1999), "What You Don't Know About Customer-Perceived Quality: The Role of Customer Expectation Distributions," *Marketing Science*, 18 (Winter), 77–95.

Schwarz, Norbert and Gerald L. Clore (1983), "Mood, Misattribution, and Judgments of Well Being: Informative and Directive Functions of Affective States," *Journal of Personality and Social Psychology*, 45 (3), 513–23.

Shih, Chuan-Fong and Alladi Venkatesh (2004), "Beyond Adoption: Development and Application of a Use-Diffusion Model," *Journal of Marketing*, 68 (January), 59–72.

Smith, Craig A. and Phoebe C. Ellsworth (1985), "Patterns of Cognitive Appraisal in Emotion," *Journal of Personality and Social Psychology*, 48 (April), 813–38.

Thompson, Debora Viana, Rebecca W. Hamilton, and Roland T. Rust (2005), "Feature Fatigue: When Product Capabilities Become Too Much of a Good Thing," *Journal of Marketing Research*, 42 (November), 431–42.

Tsiros, Michael and Vikas Mittal (2000), "Regret: A Model of Its Antecedents and Consequences in Decision Making," *Journal of Consumer Research*, 26 (March), 401–418.

Voss, Glenn B., A. Parasuraman, and Dhruv Grewal (1998), "The Role of Price, Performance, and Expectations in Determining Satisfaction in Service Exchanges," *Journal of Marketing*, 62 (October), 46–62.

Westbrook, Robert A. and Richard L. Oliver (1991), "The Dimensionality of Consumption Emotion Patterns and Consumer Satisfaction," *Journal of Consumer Research*, 18 (June), 84–91.

Wood, Stacy L. and John G. Lynch Jr. (2002), "Prior Knowledge and Complacency in New Product Learning," *Journal of Consumer Research*, 29 (December), 416–26.

BOOKS FROM THE MARKETING SCIENCE INSTITUTE

Essential Readings in Marketing

NEW

Edited by Leigh McAlister, University of Texas at Austin, Ruth N. Bolton, Arizona State University, and Ross Rizley, Marketing Science Institute

For readers seeking to understand the critical milestones of marketing thought, this collection will be an invaluable resource. Abstracts of over 200 award-winning papers are organized by research topic, with chapter introductions providing context and insight.

MARKETING SCIENCE INSTITUTE, APRIL 2006, 196 PAGES, $30

> **"An excellent and much-needed source of award-winning research in marketing for doctoral students, scholars, and business libraries."**
>
> — V. "SEENU" SRINIVASAN, Adams Distinguished Professor of Management
> Stanford University Graduate School of Business

Assessing Marketing Strategy Performance

Edited by Christine Moorman, Duke University and Donald R. Lehmann, Columbia University

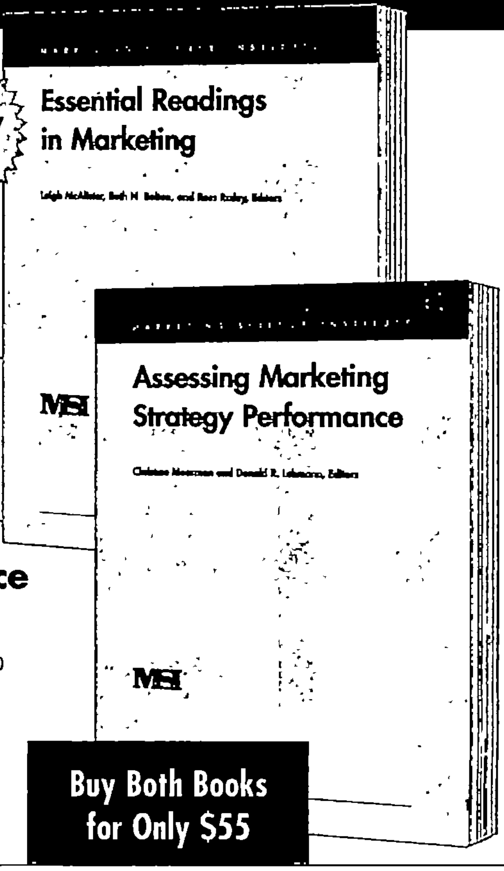
In 13 chapters, authored by leading marketing scholars, this book offers a powerful set of tools to help researchers investigate how marketing influences firm performance.

MARKETING SCIENCE INSTITUTE, AUGUST 2004, 292 PAGES, $35

To order:
www.msi.org/links/bkoffer2006.cfm

MSI

Buy Both Books for Only $55


Essential Readings in Marketing
Assessing Marketing Strategy Performance
MSI
MSI

Thorsten Hennig-Thurau, Markus Groth, Michael Paul, & Dwayne D. Gremler

Are All Smiles Created Equal? How Emotional Contagion and Emotional Labor Affect Service Relationships

In this study, the authors examine the effects of two facets of employee emotions on customers' assessments of service encounters. Drawing on emotional contagion and emotional labor theories, they investigate the influence of the extent of service employees' display of positive emotions and the authenticity of their emotional labor display on customers' emotional states and, subsequently, on customers' assessments of the service interaction and their relationship with the service provider. To test the study hypotheses, 223 consumers participated in a simulated service encounter in which actors played the roles of service employees. In a 2 × 2 factorial design, the employees varied both the extent of their smiling behavior and their emotional labor display by engaging in surface or deep acting. The results show that the authenticity of employees' emotional labor display directly affects customers' emotional states. However, contrary to expectations, the extent of employee smiling does not influence customer emotions, providing no support for the existence of primitive emotional contagion in service interactions. Furthermore, employee emotions exert an influence on customer outcomes that are of interest to marketers.

In general, the interaction between service employees and customers is considered an essential part of both customers' assessments of service quality and their relationship with the service provider (Bitner 1990; Gwinner, Gremler, and Bitner 1998; Parasuraman, Zeithaml, and Berry 1985). Despite the considerable amount of empirical research on service relationships and customer assessments of service quality, several aspects of the service interaction have remained unexplored. An area of particular interest is the role of emotions in service encounters. Although the notion of having a friendly service staff and providing "service with a smile" has become a generally unquestioned mantra for service firms, empirical research about how employees' emotional states affect customers and their assessments of service encounters has emerged only in recent years.

Two research streams that address the role of emotions in service encounters involve emotional contagion (Hatfield, Cacioppo, and Rapson 1994) and emotional labor (Hochschild 1983). "Emotional contagion" is defined as the flow of emotions from one person to another, with the receiver "catching" the emotions that the sender displays (Schoenewolf 1990). In the context of service interactions, emotional contagion creates a ripple effect of emotions from service employees to customers (Pugh 2001; Tsai and Huang 2002; Verbeke 1997). In other words, employees who smile at customers may be contagious, in that they change the affective state of customers and thus influence customers' perceptions and evaluations of the service encounter. "Emotional labor" refers to service employees' display of expected emotions as a self-regulatory process (Hochschild 1983). When displaying expected emotions to customers, employees can choose between two acting strategies, surface or deep acting, which differ mainly in their extent of authenticity (Grandey 2003).

The purposes of this study are twofold. First, we attempt to extend marketing theory by developing and testing a model of how employee emotions affect customers. Building on emotional contagion and emotional labor theories, we develop a model that enables us to test the differential effects of two facets of employee emotions—employee smiling behavior and authenticity of the emotional labor display—on changes in customer affect in a service setting. We also examine key customer consequences of interest to service firms, including the impact of employee emotions on customer satisfaction, customer–employee rapport, and customers' future loyalty intentions, all of which constitute previously untested relationships.

Thorsten Hennig-Thurau is Professor of Marketing and Media Research (e-mail: tht@medien.uni-weimar.de), and Michael Paul is a doctoral student (e-mail: michael.paul@medien.uni-weimar.de), Department of Marketing and Media Research, College of Media, Bauhaus-University of Weimar, Germany. Markus Groth is a senior lecturer, Department of Organizational Behavior, Australian Graduate School of Management, University of New South Wales (e-mail: markusg@agsm.edu.au). Dwayne D. Gremler is Associate Professor of Marketing, Department of Marketing, College of Business Administration, Bowling Green State University (e-mail: gremler@bgsu.edu). The authors thank the students involved in the data collection for their support, especially Johanna Pauge, Farnoush Pourebrahimzadeh, and Vincent Charles for their acting performances. They also thank the three anonymous *JM* reviewers, Alicia Grandey, Doug Pugh, and Anna Mattila for their helpful and constructive comments on previous versions of this article. They are grateful to CineStar and Concorde Home Entertainment for contributing the incentives for the participants in the study.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 70 (July 2006), 58–73

A second contribution of this study is our use of an experimental research design that enables us to assess the cause–effect nature of emotional contagion processes more precisely. The extant services literature on emotional contagion has not answered the question whether customers really catch employee emotions. Responding to Pugh's (2001, p. 1026) statement that "in the absence of a true experimental design, it is not possible to conclude that contagion alone is responsible [for the variation in customer consequences]," our research establishes such causal relationships and fills an important gap in the customer–employee emotions literature.

We first review the concepts that are central to our study, namely, emotional contagion, emotional labor, and customer consequences of service interactions, and then we present a conceptual model and discuss the relationships among the model elements. Next, we describe the design of our experiment, which enables us to examine the relationships of the conceptual model using a two-way repeated measures analysis of variance (ANOVA) and partial least squares (PLS) structural equation modeling. Finally, we present the findings and their implications for services research and management.

The Role of Emotions in Service Interactions

Emotional Contagion

Research on emotional contagion attempts to explain how emotions are transmitted among people in social interactions and how "catching" another person's emotions affects the dynamics of the social interaction. Emotional contagion can occur at both subconscious and conscious levels (Barsade 2002). That is, the process of emotional contagion can be attributed to people's "tendency to automatically mimic and synchronize facial expressions, vocalizations, and movements with those of another person and, consequently, to converge emotionally" (Hatfield, Cacioppo, and Rapson 1994, p. 5) and to more conscious social comparison processes between people (Barsade 2002).

With "primitive emotional contagion" (Hatfield, Cacioppo, and Rapson 1994), the transfer of emotions from one person to another is the result of the receiver's unconscious, emotive processes. This type of emotional contagion is driven by a two-step mimicry process, in which a person (1) spontaneously imitates another person's facial expressions and other nonverbal cues, which (2) leads the person to experience the corresponding emotions through physiological links. Although the person feels the emotions that result from mimicry, the processes that lead to this emotion are often "subconscious and automatic" (Barsade 2002, p. 648). As a consequence, emotional contagion theories suggest that primitive emotional contagion is spurred by the extent to which the sender displays emotions; a greater emotional display by the sender results in higher levels of emotional contagion in the receiver.

In contrast, "conscious emotional contagion" is based on social comparison processes in which people actively search for emotions as a type of social information (Salancik and Pfeffer 1978). This search activity is considered a fundamental human behavior, which grows particularly strong in situations perceived as ambiguous (Gump and Kulik 1997). Specifically, conscious emotional contagion theory argues that people compare their mood with another person's mood and adopt the sender's emotive level when it appears appropriate (Barsade 2002).[1] For example, in the absence of other social information, people visiting an attorney for the first time can be expected to observe the attorney's emotional display and then to adopt his or her emotions as a result of their desire to search for social information and reduce perceived ambiguity. Unlike primitive emotional contagion, conscious emotional contagion is determined less by the extent to which the sender displays emotions during an interaction (e.g., frequency of smiling) and more by the authenticity with which the emotions are displayed (e.g., genuineness of a smile). When the receiver perceives the sender's emotional display as fake or disingenuous, he or she will not interpret the emotional display as adequate for reducing perceived ambiguity, so conscious emotional contagion is less likely to occur.

Emotional Labor

The concept of emotional labor, which we consider a potential driver of customers' emotional states and subsequent assessments of service interactions, refers to the "effort, planning, and control needed to express organizationally desired emotions during interpersonal transactions" (Morris and Feldman 1996, p. 987). Recent management literature has considered emotional labor in an effort to better understand how service organizations can manage employees' positive displays to customers. Furthermore, it is linked to the existence of either explicit or implicit organizational display rules (Rafaeli and Sutton 1987) that define which emotions employees are expected to display and which they should suppress in the course of interacting with customers.

In general, service employees are expected to align their displayed emotions with organizationally desired emotions through their choice of emotional labor strategies (Hochschild 1983). With regard to specific emotional labor strategies, scholars have drawn on Hochschild's (1983) distinction between surface acting and deep acting as the primary framework for service employees. In "surface acting," an employee tries to change only his or her outward behavior to exhibit the required emotions. Thus, surface acting refers to the act of displaying an emotion that is not felt and could involve both suppression of felt emotions and faking of unfelt emotions. For example, when dealing with an angry or annoying customer, an employee may simply put on a smile and pretend to be cheerful and friendly without actually feeling the emotions. In other words, surface acting constitutes the expression of feigned emotions and lacks authenticity (Grandey 2003). With "deep acting," however, employees express expected (or required) emotions by attempting to create these emotions within themselves. This strategy is similar to the method-acting technique developed by Russian director Constantin Stanislavski (1965), in which actors are taught to create self-induced true emotions by using their emotional memory and recalling prior experiences and emotions (Hochschild 1983). As an example of deep acting in the context of service delivery, Hochschild reports on flight attendants who are trained to deal with angry passengers by thinking of them as frightened, first-time fliers. This process enables the flight attendants to change their inner feelings toward the customer from annoyance to pity and sympathy. Thus, the distinction between surface acting and deep acting is conceptually aligned with the "service-as-theater" metaphor, which postulates the interaction between service employees and customers as a dramaturgical interaction in which actors (i.e., employees) perform (i.e., provide service) on stage (i.e., the service environment) in front of an audience (i.e., customers) (Grove and Fisk 1992).

---

[1] This type of emotional contagion is sometimes referred to as "emotional comparison" in the literature (e.g., Bartel and Saavedra 2000)

Consequences of Service Interactions

The display of employee emotions and the resultant customer emotions likely affect various outcomes of interest to marketing managers. We focus on three major customer consequences of service interactions that are considered particularly relevant to service companies: customer satisfaction, customer–employee rapport, and customers' future loyalty intentions.

"Customer satisfaction" is widely regarded as the cognitive assessment of a customer's emotional experience (Hunt 1993). As Oliver (1981) discusses, satisfaction is consumption specific; that is, it is related to a single consumption experience. This transaction-related characteristic is often considered the main difference between satisfaction and similar evaluative concepts, such as consumer attitude and perceived service quality, which are regularly modeled as overall constructs or general evaluations of a service and are unrelated to a specific consumption episode (Hennig-Thurau and Klee 1997).

"Customer–employee rapport" is "a customer's perception of having an enjoyable interaction with a service provider employee, characterized by a personal connection between the two interactants" (Gremler and Gwinner 2000, p. 92). As a relational concept applicable to service settings, rapport depends on one or more interactions between employees and customers. Similar to satisfaction, customer–employee rapport represents a customer's cognitive assessment of an affective state. It also is important to stress that rapport can be cultivated through a single service interaction and does not depend on a shared long-term history. Customer–employee rapport has been identified as a salient issue for service organizations because rapport exerts a strong influence on customer perceptions of service delivery and service organizations (DeWitt and Brady 2003; Gremler and Gwinner 2000).

Finally, customers' "future loyalty intentions" constitute a central component of service loyalty (Oliver 1997), which itself is defined in various ways. Most definitions focus on a customer's willingness to visit a particular firm again because of his or her positive emotions and cognitions (Oliver 1999). In empirical research, future intentions represent a frequently studied component of loyalty, when loyalty is defined and/or measured as future intentions (e.g., Fornell 1992; Rust and Zahorik 1993; Zeithaml, Berry, and Parasuraman 1996). We emphasize future loyalty intentions because we believe that, in general, the future behavior of customers is of more interest to service marketers than current consumer attitudes (i.e., attitudinal loyalty) and/or prior behavior.

A Conceptual Model of the Impact of Employee Emotions on Customers in Service Interactions

The impact of different facets of employees' emotional displays on customer emotions and their consequences for customers constitute the focus of this research. In Figure 1, we show the conceptual model and the specific hypotheses tested herein. In the following section, we discuss the proposed relationships in detail.

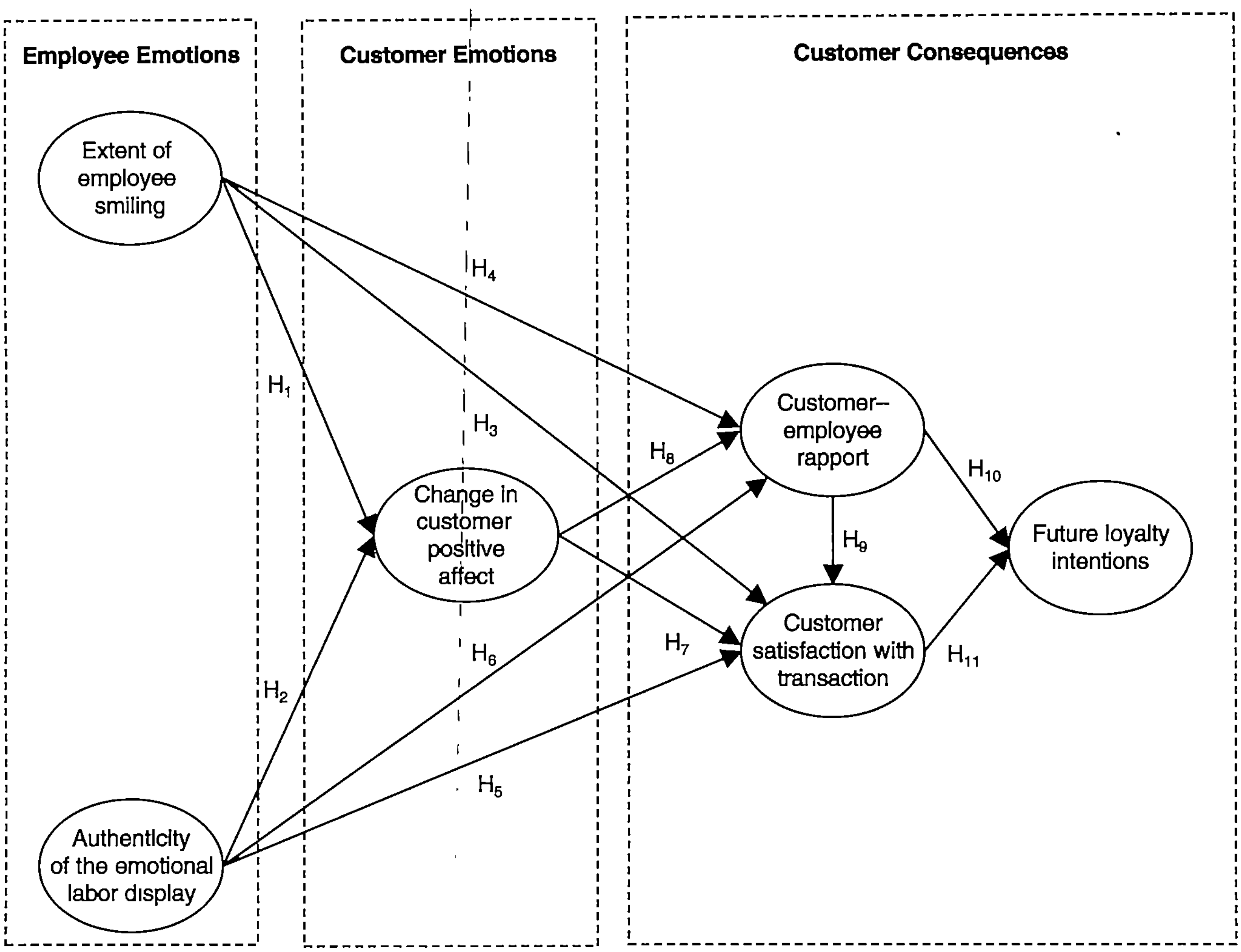
Impact of Employee Emotions on Customer Emotions

Drawing on emotional contagion and emotional labor theories, we contend that in interpersonal interactions between employees and customers, the likelihood that an employee's emotions affect customer emotions is facilitated by two key variables: the extent of an employee's smiling and the authenticity of his or her emotional labor display. Although prior research has found that employees who deliver "service with a smile" increase a customer's service experience (e.g., Rafaeli and Sutton 1989; Tsai and Huang 2002), little is known about how or why an employee's display of emotions is related to customer consequences. Emotional contagion theory (Hatfield, Cacioppo, and Rapson 1994) suggests that people's expression of positive emotions facilitates a corresponding emotional state in others. McHugo and colleagues (1985) demonstrate that exposure to images of smiling faces produces corresponding observed and self-reported emotions in study participants. Furthermore, as we discussed previously, primitive (i.e., subconscious) emotional contagion suggests that the extent of a service employee's positive emotional display (e.g., amount of smiling) is the key driver of emotional contagion (Hatfield, Cacioppo, and Rapson 1994). Thus, on the basis of the assumption that customers perceive employee smiling, we argue that if an employee increases his or her amount of smiling, customers are more likely to mimic these facial expressions subconsciously during the encounter, thus altering their own emotional state. We use the term "customer positive affect" for this positive emotional state (DeWitt and Liu 2002; Watson and Tellegen 1985).

> H_1: High amounts of employee smiling lead to a greater increase in customer positive affect than do low amounts of employee smiling.

We also expect that the authenticity of the service employee's emotional labor display influences the customer's emotional state. As we indicated previously, deep acting and surface acting are alternative strategies that

FIGURE 1
A Conceptual Model of Employee and Customer Emotions in Service Interactions


Employee Emotions
Customer Emotions
Customer Consequences
Extent of employee smiling
Authenticity of the emotional labor display
Change in customer positive affect
Customer–employee rapport
Customer satisfaction with transaction
Future loyalty intentions
H1
H2
H3
H4
H5
H6
H7
H8
H9
H10
H11

employees use to ensure that their emotional display toward customers conforms to organizational display rules. A key difference between these two emotional labor strategies is the degree of authenticity of the emotional display (Brotheridge and Grandey 2002; Hochschild 1983; Kruml and Geddes 2000). We argue that a high level of authenticity of the employee's emotional labor display, a main characteristic of deep acting, triggers positive emotions within customers due to their preference for being treated in an honest and authentic way. In contrast, we do not expect such positive customer emotions when the employee's emotional display is inauthentic, which is a main characteristic of surface acting. Evidence from the social psychological literature on authentic (i.e., Duchenne) smiles and emotion recognition supports this view. For example, Ekman and colleagues (Ekman 1992; Ekman, Davidson, and Friesen 1990; Ekman and Friesen 1982) show that authentic smiles stimulate more positive emotional reactions by respondents than do "faked" smiles. It is argued that this distinction in reactions is due to different neurological bases, in that observers often respond more positively to some subtle facial cues associated with authentic emotional displays, including the symmetry of the smile or the activation of certain muscle groups around the eyes (Ekman and Friesen 1982; Ekman, Friesen, and O'Sullivan 1988). Thus, customer reactions to authentic emotional displays are likely to be more positive than those to inauthentic displays.

This argument receives additional support from conscious emotional contagion theory (Barsade 2002; Bartel and Saavedra 2000), according to which the lack of authenticity associated with employee surface acting makes it less likely that customers can reduce their service-related ambiguity through the adoption of the employee's emotions (Grandey et al. 2005; Kruml and Geddes 2000). In the case of deep acting, however, employees summon true and genuine emotions from within and thus display positive emotions to customers that are authentic, enabling those customers to adopt the employee's emotions consciously—that is, to use the employee's displayed emotions as a type of social information to reduce the ambiguity associated with the service experience. In other words, the change in customer positive affect should be greater when employees' emotional labor displays are authentic (i.e., engage in deep acting) than when they are inauthentic (i.e., surface acting):

H_2: A high level of authenticity of the employee's emotional labor display (i.e., deep acting) leads to a greater increase in customer positive affect than does a low level of authenticity of the emotional labor display (i.e., surface acting).

Customer Consequences of Employee Emotions

In addition to the proposed effects on customer positive affect, we also expect that the independent variables (i.e., extent of employee smiling and authenticity of the emotional labor display) directly affect both customer satisfaction with the transaction and customer–employee rapport. Regarding customer satisfaction, customers often interpret an employee's emotional display as part of the service itself (i.e., the notion of service as theater; Grove and Fisk 1992), which suggests that they hold expectations about the display of positive emotions (i.e., smiling) that influence their level of satisfaction (Tsai 2001). Regarding customer–employee rapport, an employee's smiling may be an antecedent of rapport, in that it increases the receiver's enjoyment of the personal interaction (Gillis, Bernieri, and Wooten 1995; Tickle-Degnen and Rosenthal 1990).

H_3: High amounts of employee smiling lead to a greater increase in customer satisfaction with the transaction than do low amounts of employee smiling.

H_4: High amounts of employee smiling lead to a greater increase in customer–employee rapport than do low amounts of employee smiling.

Similarly, we expect that the provision of authentic emotions as part of employee deep acting leads to greater customer satisfaction and rapport than does an inauthentic emotional labor display (i.e., surface acting). Grandey (2003) finds that employees' use of deep acting leads to higher ratings of service delivery than does the use of surface acting, and Grandey and colleagues (2005) report that customer satisfaction is higher when customers perceive employee behavior as authentic. Grandey (2003) also provides support for the direct impact of the type of emotional labor strategy on customer-perceived rapport. In her study, employees' deep acting is related to perceptions of friendliness and warmth, both of which are considered characteristics of rapport (Gremler and Gwinner 2000).

H_5: A high level of authenticity of the employee's emotional labor display (i.e., deep acting) leads to a greater increase in customer satisfaction with the transaction than does a low level of authenticity of the employee's emotional labor display (i.e., surface acting).

H_6: A high level of authenticity of the employee's emotional labor display (i.e., deep acting) leads to a greater increase in customer–employee rapport than does a low level of authenticity of the employee's emotional labor display (i.e., surface acting).

Relationships Between Customer Emotions and Customer Consequences

Theoretical support for the impact of a customer's emotional state on his or her satisfaction comes from service research, which suggests that customer satisfaction with a service encounter is strongly influenced by customer emotions (Oliver 1997). In particular, when customers assess a specific consumption experience (i.e., form their level of satisfaction with the service), they draw strongly on their current emotional state and ask themselves questions such as "How do I feel about it?" Thus, a change in a customer's emotions due to an employee's emotional display should influence the customer's satisfaction. This impact is consistent with the affect-as-information model from social psychology (Schwartz and Clore 1988), which suggests that people rely on their moods as information cues when they make evaluative judgments. Consequently, a change in customer emotions facilitated by an employee's emotional display should lead to a change in customer satisfaction, such that an increase in customer positive affect results in higher levels of satisfaction.

H_7: An increase in customer positive affect has a positive effect on customer satisfaction with the transaction.

We further hypothesize that an increase in customer positive affect as a result of employee emotions will influence customers' perceptions of their rapport with the employee. Specifically, when a customer is "infected" by an employee's positive emotions as a result of a service interaction, the customer likely will enjoy the interaction with the service employee to a greater extent. Because an enjoyable interaction is a key characteristic of customer–employee rapport (Gremler and Gwinner 2000), an increase in customers' positive emotions should ultimately lead to higher levels of rapport.

H_8: An increase in customer positive affect has a positive effect on customer–employee rapport.

The services literature also indicates relationships among customer–employee rapport, customer satisfaction with the transaction, and future loyalty intentions. Specifically, we argue that rapport is positively related to customer satisfaction because an enjoyable interaction with a high degree of customer–employee rapport is usually one in which customers reveal personal information, which enables employees to customize the service offering to the customer's needs (Gremler and Gwinner 2000); both of these elements are considered integral parts of customer satisfaction. Furthermore, we expect that customer satisfaction is positively related to customers' future loyalty intentions through the creation of positive attitudes toward the service provider (Yi 1990), a claim we base on both attitude theory (Fishbein and Ajzen 1975) and exit-voice theory (Hirschman 1970). Finally, because customers who personally like a service employee and have rapport with him or her can be expected to form positive expectations about a future service experience with this employee (Gremler and Gwinner 2000), we propose that customer–employee rapport has a positive relationship to customers' future loyalty intentions.

H_9: Customer–employee rapport is positively related to customer satisfaction with the transaction.

H_{10}: Customer–employee rapport is positively related to customers' future loyalty intentions.

H_{11}: Customer satisfaction with the transaction is positively related to customers' future loyalty intentions.

Methodology

Participants

Participants in the experiment were undergraduate and graduate students from various non–business schools of a medium-sized university. The final sample contained data from 223 participants, 46.8% of whom were women. The mean age of the respondents was 23.5 years (SD = 3.2); ages range from 18 to 48 years. Although the use of student samples is sometimes considered a limitation in marketing research, drawing on student samples in experimental designs, especially those involving role playing, is well accepted for examining causal relationships (e.g., Barsade 2002; Surprenant and Solomon 1987).

We contacted participants through the university's Web site and course announcements; they received either a free movie ticket for a local movie theater or a free DVD movie in return for their participation. Before the actual data collection, we conducted a pretest with 32 consumers to assess the adequacy of the study design and the experimental manipulations and to ensure that the psychometric properties of our study measures were adequate.

Procedure

General research design. We informed all participants that a new business model for a personalized movie rental service was being tested at the university. The service was positioned as a new "movie consulting service" for which customers would answer a series of questions about their movie-viewing habits and preferences in a one-to-one consultation with a service employee (unbeknownst to the subject, this employee was a trained actor). Based on their answers, customers would then receive personalized advice for potential movie rentals from the employee. Throughout the duration of the study, professionally designed posters that promoted this new service appeared in various locations at the university. The university also set aside specific facilities to serve as the movie consulting service location for the duration of this study. The layout and interior decoration of these facilities resembled those of actual video rental services (e.g., a service counter, shelves containing movies, movie posters, and life-sized figures announcing current and upcoming movies). Donations of actual DVD and VHS movies and various poster displays from local video rental stores and movie theaters decorated the facilities and thus increased the realism of the servicescape (see Figure 2).

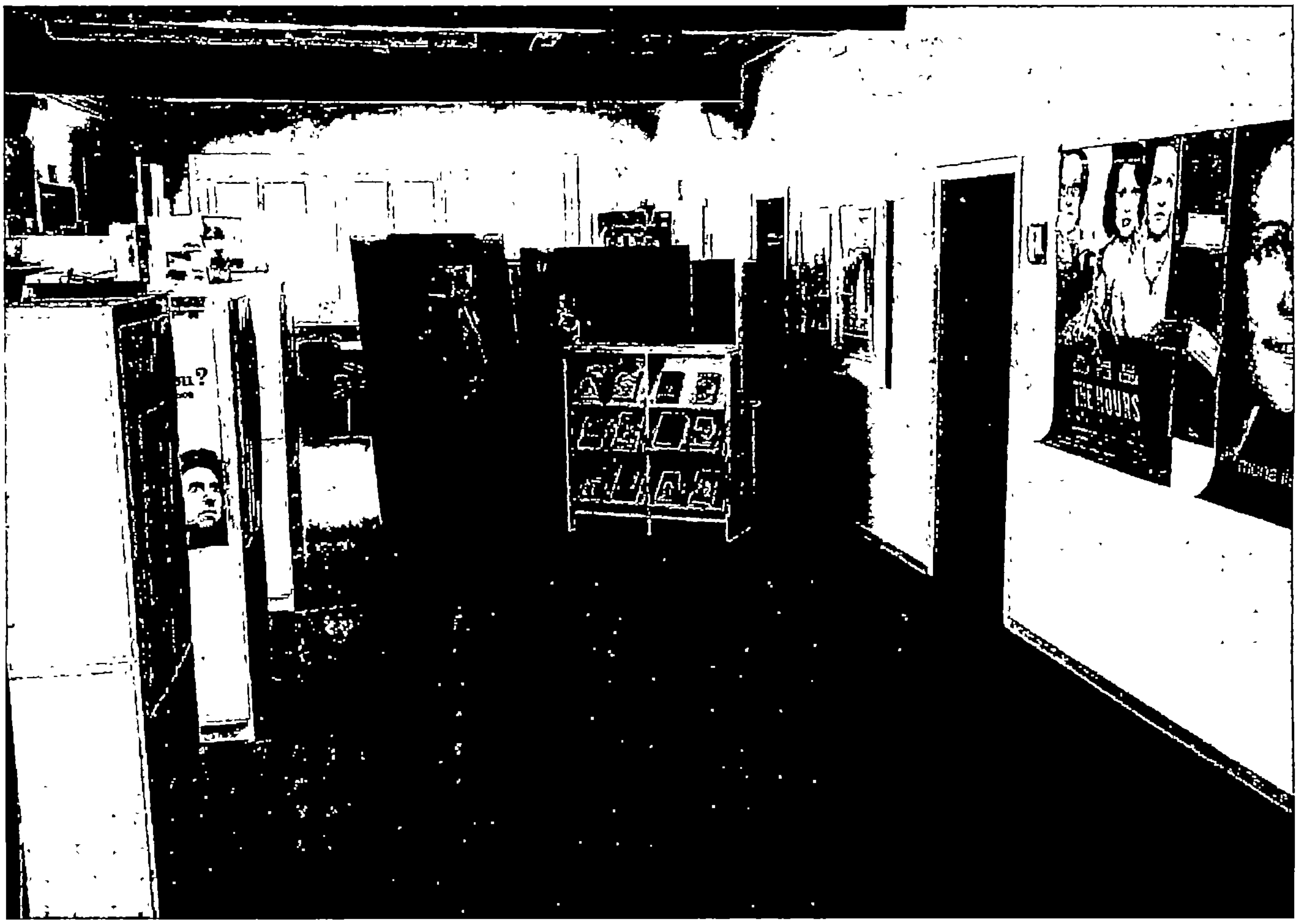
Experimental manipulations. Using a 2 × 2 between-subjects factorial design, we manipulated the extent of employee smiling (high versus low) and the authenticity of the emotional labor display (high authenticity/deep acting versus low authenticity/surface acting). We randomly assigned participants to one of the four experimental conditions, high smiling/deep acting (n = 58), high smiling/surface acting (n = 52), low smiling/deep acting (n = 55), and low smiling/surface acting (n = 58).

Dramaturgy. We used appointments to schedule one participant at a time, which limited waiting time and eliminated potential interactions among the participants. On arrival, each participant filled out a preencounter questionnaire and then immediately entered the video consulting store to begin the service encounter. During the service encounter, participants interacted with the service employee, who engaged them in a conversation by asking a series of specific questions about movie preferences and viewing habits. The service encounter concluded with the employee recommending a specific movie on DVD or VHS, which he or she then handed to the customer. In general, these service encounters lasted between 5 and 10 minutes (M = 7.5 minutes).

Actors. The service employees were three trained student actors (two female, one male) whom we specifically recruited for this study and paid for their time. We hired these actors on the basis of their experience in service work, their interest in acting, and their knowledge of motion pictures. Initially, we narrowed the pool of potential candidates down to six people, all of whom then performed in a series of auditions and rehearsals and participated in the pretest of the experiment. On the basis of their performances in these tests, we selected the final three actors and trained them over a period of six weeks.

Furthermore, we developed a role description for each of the four conditions. All roles required the actors to behave in a customer-oriented manner and avoid the display of negative emotions. For the extent of smiling manipulation, we trained the employees to smile frequently when displaying a high level of smiling and to minimize their smiling when displaying low-level smiling behavior. In line with previous research, we defined smiling as a noticeable upward twist of the employee's lips (Pugh 2001; Sutton and Rafaeli 1988). For the authenticity of the emotional labor display manipulation, we trained the actors in accordance with the characteristics of surface and deep acting, as described in the literature. They read some pertinent introductory literature (e.g., Grandey 2003; Hochschild 1983), and two of the authors then discussed the concepts with the actors for a total of ten hours. Mainly, we drew from the concept of method acting by showing movie excerpts and using teaching techniques associated with method acting. For the surface acting role (i.e., low authenticity of the emotional labor display), the employee was instructed to adapt only his or her outward behavior to the customer's needs but not his or her inner feelings. For deep acting (i.e., high authenticity of the emotional labor display), the employee was instructed to create the appropriate emotions within him- or herself. To teach these actors the skills to do so effectively, we employed specific exercises used in Stanislavski's (1965) acting technique, which are commonly taught to acting students as part of their basic training. These exercises target the development of specific skills for using emotion memories (i.e., remembering and reexperiencing specific emotions) to evoke the required emotions. We continued the training until both the actors and the researchers were satisfied that the actors had mastered a sufficient skill set to enable them to act out the specific roles required for the experiment.

FIGURE 2
Design of the Video Consulting Store



Finally, we developed a standardized service script that consisted of a list of questions about customers' movie-related preferences to ensure that all service encounters would be identical in nature in all respects other than the experimental manipulations (for further details about the service script, see Appendix A). Data collection took place over a period of seven days. Each actor played only one role per day over a time of approximately two-and-a-half hours. We randomly determined the roles in advance, and each actor performed for roughly equal amounts of time in each of the four roles.

Measures

Preencounter questionnaire. Study participants received a questionnaire just before they entered the video rental store, which included four items from Brief and colleagues' (1988) job affect scale to measure preencounter customer positive affect (see also Burke et al. 1989). We dropped two of the original six items (strong and active) because we deemed them to be inappropriate for use by customer respondents. In addition, to conceal the true nature of the study, the preencounter questionnaire contained various filler items about preferences and consumption behaviors for motion pictures.

Postencounter questionnaire. Participants completed a postencounter questionnaire immediately after the end of the service encounter; it included measures of post-encounter customer positive affect, customer–employee rapport, customer satisfaction with the transaction, and future loyalty intentions, as well as demographic variables. We assessed postencounter customer positive affect with the same items used in the preencounter survey and calculated any changes as postencounter customer positive affect less preencounter customer positive affect for each of the four items.[2] To measure rapport, we used four items from the scale developed by Gremler and Gwinner (2000); we excluded items that implicitly suggest a long-term relationship between the customer and the employee. For customer satisfaction with the transaction, we used five items from the work of Churchill and Surprenant (1982) and Bitner and Hubbert (1994). Finally, we measured future loyalty intentions with four items, three of which we adapted from the work of Zeithaml, Berry, and Parasuraman (1996) and one from the work of Taylor and Baker (1994). We measured all items using seven-point scales, in which higher numbers indicated greater agreement with the construct (for a list of all study items, see Appendix B).

---

[2] There is an ongoing discussion about the problems associated with the use of difference scores (e.g., Peter, Churchill, and Brown 1993) These issues predominantly address the integration of two measures (e.g., service expectations and service performance evaluation) but are less relevant when the same construct is measured at two points in time. Two common concerns about difference scores are that they lack variability and do not differ from the initial level (Tisak and Smith 1994). However, these issues do not

Results

Manipulation Checks

We employed several methods to ensure the validity of our study manipulations. First, unbeknownst to customers, all service encounters were recorded with a hidden digital video camera and transmitted in real time to an adjacent room. Two trained judges (one of the authors and one graduate student) watched each interaction to judge whether it was consistent with the required experimental manipulation. Service encounters that the judges deemed to be inconsistent with the required role were dropped from further analyses. Second, because the authenticity of the emotional labor display manipulation ultimately depends on an employee's self-evaluation, after each service encounter, we asked the service employees whether they believed they had effectively portrayed the desired role. In addition, during the debriefing, one participant guessed the true nature of the study. As a result of these measures, we excluded three service encounters from further analysis, which resulted in 223 usable encounters.

Third, after completion of the study, four graduate students (two male, two female) who were not familiar with the study watched all 223 videotaped service encounters and rated each interaction on several dimensions.[3] Specifically, two of the judges separately watched all interactions and recorded (1) the frequency of employee smiles (defined as the number of times the employee noticeably smiled) during each interaction and (2) the duration of each smile (measured in seconds). We transformed these data into a smiles-per-minute measure and a measure of the percentage of time the employee smiled during each encounter, which enabled us to control for service encounter length. The interrater agreement, which we measured with the intraclass correlation coefficient (ICC; Shrout and Fleiss 1979), was high between the two judges for both items (frequency of smiles: ICC = .909, $p < .01$; duration of smiles: ICC = .948, $p < .01$). The ANOVA results also showed that the manipulation of the extent of employee smiling had a significant effect on the frequency of smiles per minute ($M_{\text{high extent}}$ = 3.5, $M_{\text{low extent}}$ = .4; $F(1, 220) = 386.62$, $p < .01$) and the percentage of time spent smiling ($M_{\text{high extent}}$ = 35.2%, $M_{\text{low extent}}$ = 1.8%; $F(1, 220) = 417.09$, $p < .01$), as intended. We found no significant effect for the authenticity of the emotional labor display on the frequency of smiles ($M_{\text{high authenticity}}$ = 2.0, $M_{\text{low authenticity}}$ = 1.8; $F(1, 220)$ = 1.08, not significant [n.s.]), though it influenced the percentage of time spent smiling ($M_{\text{high authenticity}}$ = 23.9%, $M_{\text{low authenticity}}$ = 12.6%; $F(1, 220) = 18.64$, $p < .01$), which suggests a potential confounding of the manipulations. As Perdue and Summers (1986, p. 323) suggest, providing evidence that the effect sizes for unintended manipulations are substantially smaller than those for the intended manipulations indicates that the experimental manipulations have worked effectively and that statistical significance "should not be of great concern." In our case, the extent of the employee smiling manipulation had a much greater effect on the percentage of time spent smiling (partial $\eta^2 = .656$) than did the authenticity of the emotional labor display manipulation (partial $\eta^2 = .078$), thus providing support for the validity of our experimental manipulation.

The other two judges also watched all 223 videotaped interactions and separately provided a global rating of the authenticity of the employee's emotional display for each interaction (on a four-point scale ranging from "not at all authentic" [1] to "very authentic" [4]). Again, interrater agreement was high between the two judges (ICC = .927, $p < .01$). The authenticity of the emotional labor display manipulation had a significant effect on authenticity ratings ($M_{\text{high authenticity}}$ = 3.8, $M_{\text{low authenticity}}$ = 1.2; $F(1, 220)$ = 2552.78, $p < .01$), as intended. The extent of employee smiling also had a significant main effect on authenticity ratings ($M_{\text{high extent}}$ = 2.7, $M_{\text{low extent}}$ = 2.3; $F(1, 220) = 4.21$, $p < .05$), but the effect size of the authenticity of the emotional labor display manipulation (partial $\eta^2 = .932$) was greater than that of the employee smiling manipulation (partial $\eta^2 = .147$), again providing support for the validity of our experimental manipulation.

Reliability and Validity Assessment

We report the means, standard deviations, and correlation coefficients of all variables in Table 1. To assess the reliability and validity of our measures, we calculated Cronbach's alpha coefficients for each construct. All alpha scores are satisfactory, with no values below .80. In addition, we conducted a confirmatory factor analysis with all multi-item measures in the model. The overall fit statistics for the four-factor model indicate that the model provides an acceptable fit to the data: $\chi^2(129, N = 223) = 333.7$, $p <$ .01; comparative fit index = .950; incremental fit index = .950; Tucker–Lewis index = .941; standardized root mean square residual = .048; and root mean square error of approximation = .086.

The average variance extracted is greater than .60 for all constructs, and the composite reliability measures are all greater than .80. We also find support for convergent validity because the t-values for all constructs are significant at $p < .01$ (Anderson and Gerbing 1988). Similarly, we

---

appear to be critical in our case. Specifically, the standard deviations of the difference scores range from 1.46 to 1.65, and the correlation between the difference score and preencounter affect (–.34), though significant, is smaller than 1.

[3] We thank the anonymous reviewers for their suggestions of how to use the videotaped encounters to provide additional manipulation checks for this study.

TABLE 1
Descriptive Statistics, Reliability Information, and Correlations

	Number of Items	M	SD	1	2	3	4	5	6
1. Customer positive affect (preencounter)	4	3.70	1.20	.812					
2. Customer positive affect (postencounter)	4	4.41	1.41	.556	.904				
3. Change in customer positive affect	4	.71	1.24	−.336	.596	.806			
4. Customer–employee rapport	4	4.91	1.66	.233	.643	.503	.967		
5. Customer satisfaction with transaction	5	4.69	1.47	.323	.771	.561	.827	.932	
6. Future loyalty intentions	4	5.12	1.56	.269	.662	.486	.696	.798	.940

Notes: N = 223 Values along the diagonal represent Cronbach's alpha internal consistency estimates All correlations are significant at $p < .01$ (two tailed).

find support for discriminant validity; the squared correlations between each pair of constructs are smaller than the average variance explained of the respective constructs (Fornell and Larcker 1981). Thus, these results indicate the acceptable reliability, convergent validity, and discriminant validity of our measures.

Two-Way Repeated Measures ANOVA

To test H_1 and H_2, we conducted a two-way repeated measures ANOVA to examine the effects of the extent of employee smiling and the authenticity of the emotional labor display on the change in customer positive affect. The within-subjects factor, customer positive affect, contains two levels: preencounter affect (Time 1) and postencounter affect (Time 2). As Tables 2 and 3 show, the extent of employee smiling has no significant main effect on customer positive affect (F(1, 217) = 1.5, n.s.), but the authenticity of the emotional labor display by employees has a significant effect (F(1, 217) = 35.2, $p < .01$, partial $\eta^2 = .139$).[4] The nature of this effect is in the expected direction: Employees who express authentic emotions by engaging in deep acting facilitate customer positive affect to a much greater extent (change in customer positive affect = 1.18, SD = 1.09) than do those who display inauthentic emotional labor by engaging in surface acting (change in customer positive affect = .25, SD = 1.22). This finding holds true not only for the overall composite score of customer affect but also for each of the four emotions (i.e., elated, peppy, enthusiastic, and excited). Thus, we find strong evidence in support of H_2 but not of H_1. Therefore, we conclude that employees' emotional displays affect customers' emotional states, but this process appears to be driven primarily by the authenticity of the emotional labor display rather than by primitive emotional contagion through employee smiling.

PLS Equation Modeling

Next, we tested all elements of our conceptual model simultaneously with PLS structural equation modeling. As a component-based method, PLS permits the use of nominal data, which we need to assess the effects of smiling and authenticity (Fornell and Bookstein 1982). Furthermore, as a distribution-free method, PLS has fewer constraints and statistical specifications than covariance-based techniques, such as LISREL.

We operationalized both the extent of employee smiling and the authenticity of the emotional labor display as dichotomous variables (1 = low extent, 2 = high extent; 1 = low authenticity/surface acting, 2 = high authenticity/deep acting, respectively). To estimate the model paths, we used PLS Graph 3.0 and estimated the inner weightings with the path method (Chin 2001). To estimate the predictive power of the model, we applied a blindfolding approach (Fornell and Bookstein 1982). This approach results in Q^2 values of .38 (change in customer positive affect), .67 (customer satisfaction), .81 (rapport), and .72 (loyalty intentions), all of which are significantly different from 0 and therefore indicate that the model has predictive power (Geisser 1974; Stone 1974). The model explains 14.8% of the variance in change in customer positive affect, 47.6% of the variance in rapport, 71.7% of the variance in customer satisfaction with the transaction, and 65.9% of the variance in future loyalty intentions, in further support of its relevance (Chin 1998). All indicator reliabilities are greater than .75, composite reliability is greater than .85 for all constructs, and the average variance explained is greater than .60 in all cases (for more detailed information, see Appendix B). To determine whether the use of a difference score measure for change in positive affect influenced the results, we also analyzed an alternative model with postencounter customer positive affect rather than change in customer positive affect, but we found no differences between the two models.[5]

The PLS results support most of our hypotheses about the inner model relationships and provide convergent validity for the findings of the two-way repeated measures ANOVA (see Table 4).[6] Specifically, our results show that the authenticity of the employee's emotional labor display

---

[4] Although we did not formally hypothesize an interaction effect, our examination indicates that both effects are largely independent with no interaction effect (F(1, 217) = .607, n.s.). We also investigate the extent to which the results are influenced by the individual employees by controlling for actors in the analysis. This did not change the results.

[5] Significance levels and the direction of the effects for all hypothesized paths are identical to those of the difference score model. The variance explanations for the postencounter model are 11.6% for postencounter customer positive affect, 60.4% for rapport, 78.4% for customer satisfaction, and 66.0% for loyalty intentions.

[6] We generated the t-values through a bootstrapping procedure with 223 resamples with 100 cases each (Fornell and Bookstein 1982).

TABLE 2
Results of the Two-Way Repeated Measures ANOVA

		Overall Customer Positive Affect		Elated		Peppy		Enthusiastic		Excited	
Independent Variable	**d.f.**	**F**	**Partial η^2**	**F**	**Partial η^2**	**F**	**Partial η^2**	**F**	**Partial η^2**	**F**	**Partial η^2**
Extent of employee smiling (SMILE)	1	1.549		2.708		1.309		.190		1.179	
Authenticity of the emotional labor display (AUTH)	2	35.175*	.139	23.565*	.098	12.943*	.057	16.185*	.070	31.388*	.128
SMILE × AUTH	2	.607		.348		.938		.002		1.049	
Error	217	(.666)		(1.193)		(1.095)		(1.010)		(1.181)	

*$p < .01$.

Notes: We calculate the overall composite score for customer positive affect as the mean value of the four emotion items. Values in parentheses are mean square terms.

TABLE 3
Means and Standard Deviations for Customer Positive Affect

		Customer Positive Affect		
		Time 1: Preencounter	**Time 2: Postencounter**	**Change in Customer Positive Affect**
Extent of employee smiling	High extent	3.72 (1.21)	4.54 (1.38)	.82 (1.26)
	Low extent	3.69 (1.20)	4.30 (1.44)	.61 (1.21)
Authenticity of the emotional labor display	High authenticity	3.71 (1.20)	4.87^a (1.28)	1.18^b (1.09)
	Low authenticity	3.70 (1.21)	3.95^a (1.40)	.25^b (1.22)

Notes: Standard deviations are shown in parentheses. Values with the same superscript are significantly different from each other at $p < .001$; all other values do not significantly differ.

has a strong and significant impact on the change in customer positive affect—high authenticity through deep acting results in a significantly greater increase in the change in customer positive affect than does low authenticity through surface acting—but the extent of employee smiling does not significantly influence the change in customer positive affect.

The extent of employees' smiling and the authenticity of their emotional labor display both have a significant, direct impact on customer–employee rapport, in support of H_4 and H_6, but they do not affect customer satisfaction, which fails to support H_3 and H_5. The PLS results also show that change in customer positive affect has a significant effect on both customer satisfaction with the transaction and customer–employee rapport, in support of H_7 and H_8, respectively. Rapport positively influences customer satisfaction (in support of H_9), and customer satisfaction positively influences future loyalty intentions (in support of H_{11}). Contrary to our expectations, rapport has no direct relationship to future loyalty intentions, leading us to reject H_{10}. However, the total effect of rapport on loyalty intentions is strong and significant as a result of the strong impact of rapport on satisfaction (for a listing of all total effects, see Table 4).

Discussion

What We Have Learned

The purpose of this study was to develop and experimentally test a conceptual model of how employee emotions influence customers in service encounters. We apply a two-way repeated measures ANOVA and PLS structural equation modeling to data collected from a sample of 223 consumers. The results of this research provide several important contributions to the literature.

First, we fill an important gap in the service marketing literature by providing empirical evidence that an employee's emotional display can trigger changes in a customer's affective state. Our study is the first to provide a direct test of emotional contagion and its related processes in service interactions, and the controlled experimental design of our study enables us to test causal effects between different facets of an employee's display of emotions and the customer's state.

Second, our findings suggest that the authenticity of the employee's emotional labor display, rather than the extent of smiling, influences the customer's emotions and perceptions. Respondents in our study who encountered authentic employees (i.e., those who engaged in deep acting) were far more likely to adopt the emotions of that employee than were those who interacted with inauthentic employees (i.e., those who engaged in surface acting). This finding contributes to the emotional labor literature in that our study is the first to provide evidence that different types of emotional labor (i.e., deep versus surface acting) differentially influence the customer experience during service encounters. Furthermore, our findings contribute to the emotional contagion literature by contradicting the dominant belief that emotional contagion in service settings is based mainly on mimicry effects (i.e., primitive contagion processes; see, e.g., Pugh 2001). Although mimicry effects occur within extremely short time frames (often less than one second) and thus might affect the receiver's emotions for a short period, our results suggest that their impact fades during the course of a service encounter. That is, primitive emotional contagion might occur in the early phases of service encounters, but it does not remain throughout the completion of the encounter, because postencounter emotions do not appear to be influenced by mimicry effects.

Third, our findings provide support for the affect-as-information theory, which links customer emotions to satisfaction (Schwartz and Clore 1988); the results show that both the customer's postencounter emotional state and change in positive affect lead to increased customer satisfaction. An increase in customer positive affect also influences customer satisfaction indirectly through customer–employee rapport. Because satisfaction is positively related to customers' future loyalty intentions, employees' emotional displays appear to play important roles in a service firm's long-term success and have significant impacts on key customer outcomes.

Fourth, we could not replicate Grandey and colleagues' (2005) results of a direct relationship between emotional labor and customer satisfaction. In our study, the two paths from the extent of employee smiling and the authenticity of the emotional labor display to customer satisfaction are not significant. Therefore, our results suggest that such effects are of an indirect nature, with customer–employee rapport being the main driver of customer satisfaction. Thus, the

TABLE 4
Path Coefficients from PLS Analysis

Hypothesis	Impact of	On	Path Coefficient	t-Value	Total Impact	Hypothesis Supported?
H_1	Extent of employee smiling	Change in customer positive affect	.070	1.081	.070	No
H_2	Authenticity of the emotional labor display	Change in customer positive affect	.375*	4.572	.375	Yes
H_3	Extent of employee smiling	Customer satisfaction with transaction	–.046	–1.081	.014	No
H_4	Extent of employee smiling	Customer–employee rapport	.194*	2 752	.217	Yes
H_5	Authenticity of the emotional labor display	Customer satisfaction with transaction	–.055	–1.100	.384	No
H_6	Authenticity of the emotional labor display	Customer–employee rapport	.453*	6.654	.574	Yes
H_7	Change in customer positive affect	Customer satisfaction with transaction	.206*	3.041	.453	Yes
H_8	Change in customer positive affect	Customer–employee rapport	.323*	5.041	.323	Yes
H_9	Customer–employee rapport	Customer satisfaction with transaction	.765*	11.380	.765	Yes
H_{10}	Customer–employee rapport	Future loyalty intentions	.109	1.342	.651	No
H_{11}	Customer satisfaction with transaction	Future loyalty intentions	.720*	7.491	.720	Yes

*$p < .05$

link between authenticity of the emotional labor display and customer satisfaction is not as direct as previously thought but instead is mediated by the extent to which customers perceive that rapport with the service employee exists. A possible explanation of this finding is that whereas the construct of rapport focuses mainly on the interaction with the employee, satisfaction also captures the customer's assessment of other aspects of the service encounter, such as the design of the servicescape (i.e., "I like how the video store looks") and outcome qualities (i.e., "I like the movie I was given by the employee").

Fifth, although we find that customer–employee rapport is a driver of satisfaction, our results also suggest that rapport does not exert a direct influence on customer loyalty. Although previous marketing research has postulated such a direct relationship (Gremler and Gwinner 2000), our study suggests that the interplay among rapport, satisfaction, and future loyalty intentions may be more complex than previously believed.

What We Still Need to Learn

Although our findings expand the extant knowledge on emotions in service encounters, we recognize several limitations that must be taken into account when generalizing our results. The design of our experiment provides insight into the transfer of employee emotions in one specific service domain—an innovative video rental service—but it remains unclear whether findings would be similar for other services. Another limitation is that we focus on positive emotions because they are most relevant (and desired) in the service delivery context. However, similar ripple effects may occur for negative emotions. For example, anger and frustration displayed by employees may transfer to customers and negatively affect their service experience. Further research should address this issue by testing whether effects similar to those we find apply equally to the flow of negative emotions in the service delivery context.

Because participants in our study engaged in a service that was unfamiliar to them, their emotional processes were completely independent of any prior service interactions. However, in many real-life services, relationships between employees and customers have been established over a series of encounters, which might influence the emotional flow in future service transactions. Another avenue for further research would be to test the cross-cultural stability of our findings. Given that cultural norms about service delivery and smiling behavior in general vary across cultures, additional research should examine the extent to which cultural variables influence the interplay of employee emotional displays and customer variables in service deliveries.

What Marketing Managers Should Do

This study illustrates that the emotions displayed by frontline employees are an important driver of the relationships between service employees and customers. The transfer of positive emotions from employees to customers can enable firms to establish high levels of customer–employee rapport, customer satisfaction, and future loyalty intentions, all key objectives of relationship marketing.

Our finding that customer affect is driven mainly by the authenticity of the emotional labor display, rather than the extent of employee smiling, has important implications for service marketers. First, it contradicts the common mantra in service organizations that service with a smile and friendly service staff inevitably pay off and serve as a competitive advantage. The commonly advocated "Smile!" policy that is currently in place in many firms may be a less useful strategy than is generally believed. Rather, our results indicate that the authenticity of the emotional display by frontline staff and the sincerity with which staff interacts with customers may be much stronger drivers of service outcomes than policies that require people to smile at any cost but do not provide them with the tools to create and display seemingly real emotions when they serve customers. Simply hiring low-paid service workers with limited qualifications and motivation and requiring them to smile at customers as part of their job description may not deliver the desired results, whereas the display of authentic feelings by service employees—facilitated by the use of deep acting—is likely to be more effective for positively influencing customer satisfaction and related service outcomes.

Second, in terms of managing frontline employees, the provision of authentic feelings is a far more challenging task than simply smiling at customers. Such strategy may require increased managerial emphasis on identifying and hiring talented and qualified frontline employees. Managers also may need to provide additional training for frontline staff that teaches them how to engage in deep acting. One such approach uses a perspective-taking technique, which puts employees in the customers' shoes and thereby increases their ability to adopt a customer's viewpoint (Parker and Axtell 2001). For example, flight attendant trainees can be taught strategies to deal with irate customers, such as imagining that something traumatic has happened in these customers' lives or by pretending that they are children (Hochschild 1983).

Third, our results confirm that the emotions customers experience during service encounters play crucial roles and directly affect the success of service relationships. Because customer emotions appear to be key drivers of rapport with employees and, ultimately, customer satisfaction and loyalty intentions, service organizations may benefit from focusing their attention on increasing positive customer emotions. This recommendation is consistent with emerging literature on customer delight (Rust and Oliver 2000), which stresses the emotional component of customers' service evaluations.

APPENDIX A
Summary of Service Script

Step 1. As soon as the customer enters the video consulting store and steps toward the service counter, the employee makes eye contact with the customer and welcomes him or her to the store.

Step 2. The employee briefly introduces the nature of the movie consulting business and its service offerings following a standardized script. The employee points out that this service is currently being tested and thus is free of charge to the customer.

Step 3. The employee asks a series of standardized questions about the customer's movie-viewing habits and preferences. These questions follow a standardized order and cover the following topics: (1) preferences for specific movie genres, (2) preferences for happy endings, (3) importance of film directors in influencing choice of movies, (4) preferences for specific film directors, (4) preferences for specific lead actors and actresses, and (5) preferences for a movie's country of origin. During this process, the employee takes brief notes while the customer answers all questions.

Step 4. Following the set of standardized questions, the employee briefly reviews his or her notes and summarizes the customer's responses.

Step 5. The employee proceeds to recommend a specific movie to the customer on the basis of the customer's responses. The employee briefly introduces the movie using a standardized script. If the customer has already seen this movie or is not interested for other reasons, the employee recommends up to two additional movies.

Step 6. When the movie choice is finalized, the employee asks the customer for his or her choice of DVD or VHS format, and the employee hands the movie to the customer. The customer is thanked for his or her business and leaves the video consulting store.

APPENDIX B
List of Items and Goodness-of-Fit Measures of PLS Analysis

Change in Customer Positive Affect (Composite Reliability [CR] = .871, Average Variance Extracted [AVE] = .627)

At this moment, I feel ...	
elated.	.794
peppy.	.760
enthusiastic.	.800
excited.	.813

Customer Satisfaction with Transaction (CR = .948, AVE = .784)

I am delighted by this service experience.	.922
This service experience really helped me to find a good movie.	.764
This service experience was a great one.	.863
I am satisfied with this specific service experience.	.939
I really liked this service experience.	.928

Customer–Employee Rapport (CR = .977, AVE = .913)

In thinking about my relationship with this person, I enjoyed interacting with this employee.	.949
This employee created a feeling of "warmth" in our relationship.	.958
This employee related well to me.	.947
I was comfortable interacting with this employee.	.969

Future Loyalty Intentions (CR = .956, AVE = .846)

The next time I need the services of a video store, I will choose this company.	.936
I will say positive things about this service provider to other people.	.922
I plan to visit this service provider in the next years.	.922
I will recommend this service provider to someone who seeks my advice.	.898

Notes: Numbers in the right-hand column are coefficients of determination

REFERENCES

Anderson, James C. and David W. Gerbing (1988), "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin*, 103 (3), 347–62.

Barsade, Sigal G. (2002), "The Ripple Effect: Emotional Contagion and Its Influence on Group Behavior," *Administrative Science Quarterly*, 47 (December), 644–75.

Bartel, Caroline A. and Richard Saavedra (2000), "The Collective Construction of Work Group Moods," *Administrative Science Quarterly*, 45 (June), 197–231.

Bitner, Mary Jo (1990), "Evaluating Service Encounters: The Effects of Physical Surroundings and Employee Responses," *Journal of Marketing*, 54 (April), 69–82.

——— and Amy R. Hubbert (1994), "Encounter Satisfaction Versus Overall Satisfaction Versus Quality: The Customer's Voice," in *Service Quality: New Directions in Theory and Practice*, Roland T. Rust and Richard L. Oliver, eds. Thousand Oaks, CA. Sage Publications, 72–94.

Brief, Arthur P., Michael J. Burke, Jennifer M George, Brian S. Robinson, and Jane Webster (1988), "Should Negative Affectivity Remain an Unmeasured Variable in the Study of Job Stress?" *Journal of Applied Psychology*, 73 (May), 193–98.

Brotheridge, Céleste M. and Alicia A. Grandey (2002), "Emotional Labor and Burnout: Comparing Two Perspectives of 'People Work,'" *Journal of Vocational Behavior*, 60 (February), 17–39.

Burke, Michael J., Arthur P. Brief, Jennifer M. George, Loriann Roberson, and Jane Webster (1989), "Measuring Affect at Work: Confirmatory Analyses of Competing Mood Structures with Conceptual Linkage to Cortical Regulatory Systems," *Journal of Personality and Social Psychology*, 57 (6), 1091–1102.

Chin, Wynne W. (1998), "The Partial Least Square Approach to Structural Equations Modeling," in *Modern Methods for Business Research*, George A. Marcoulides, ed. Mahwah, NJ: Lawrence Erlbaum Associates, 295–336.

——— (2001), *PLS-Graph User's Guide: Version 3.0.* Houston: Soft Modeling Inc.

Churchill, Gilbert A. and Carol Surprenant (1982), "An Investigation into the Determinants of Customer Satisfaction," *Journal of Marketing Research*, 19 (November), 491–504.

DeWitt, Tom and Michael K. Brady (2003), "Rethinking Service Recovery Strategies," *Journal of Service Research*, 6 (November), 193–207.

——— and Yongmei Liu (2002), "The Customer Orientation-Loyalty Model: The Role of Emotional Contagion and Rapport in the Service Encounter," in *AMA Summer Educators' Conference: Enhancing Knowledge Development in Marketing*, William J. Kehoe and John H. Lindgren Jr., eds. Chicago: American Marketing Association, 321.

Ekman, Paul (1992), "Facial Expressions of Emotions: New Findings, New Questions," *Psychological Science*, 3 (January), 34–38.

———, Richard J. Davidson, and Wallace V. Friesen (1990), "The Duchenne Smile: Emotional Expression and Brain Psychology II," *Journal of Personality and Social Psychology*, 58 (2), 342–58.

——— and Wallace V. Friesen (1982), "Felt, False, and Miserable Smiles," *Journal of Nonverbal Behavior*, 6 (4), 238–52.

———, ———, and Maureen O'Sullivan (1988), "Smiles When Lying," *Journal of Personality and Social Psychology*, 54 (3), 414–20.

Fishbein, Martin and Icek Ajzen (1975), *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.

Fornell, Claes (1992), "A National Customer Satisfaction Barometer: The Swedish Experience," *Journal of Marketing*, 56 (January), 6–21.

——— and Fred L. Bookstein (1982), "Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory," *Journal of Marketing Research*, 19 (November), 440–52.

——— and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39–50.

Geisser, Seymour (1974), "A Predictive Approach to the Random Effects Model," *Biometrika*, 61 (1), 101–107.

Gillis, John S., Frank J. Bernieri, and Eric Wooten (1995), "The Effects of Stimulus Medium and Feedback on the Judgment of Rapport," *Organizational Behavior and Human Decision Processes*, 63 (July), 33–46.

Grandey, Alicia A. (2003), "When 'The Show Must Go On': Surface and Deep Acting as Determinants of Emotional Exhaustion and Peer-Rated Service Delivery," *Academy of Management Journal*, 46 (February), 86–96.

———, Glenda M. Fisk, Anna S. Mattila, Karen J. Jansen, and Lori A. Sideman (2005), "Is 'Service With a Smile' Enough? Authenticity of Positive Displays During Service Encounters," *Organizational Behavior and Human Decision Processes*, 96 (January), 38–55.

Gremler, Dwayne D and Kevin P. Gwinner (2000), "Customer-Employee Rapport in Service Relationships," *Journal of Service Research*, 3 (August), 82–104.

Grove, Stephen J. and Raymond P. Fisk (1992), "The Service Experience as Theater," in *Advances in Consumer Research*, Vol. 19, John F. Sherry and Brian Sternthal, eds. Provo, UT: Association of Consumer Research, 455–61

Gump, Brooks B. and James A. Kulik (1997), "Stress, Affiliation, and Emotional Contagion," *Journal of Personality and Social Psychology*, 72 (2), 305–319.

Gwinner, Kevin P, Dwayne D. Gremler, and Mary Jo Bitner (1998), "Relational Benefits in Services Industries: The Customer's Perspective," *Journal of the Academy of Marketing Science*, 26 (Spring), 101–114.

Hatfield, Elaine, John T. Cacioppo, and Richard L. Rapson (1994), *Emotional Contagion*. Cambridge, UK: Cambridge University Press.

Hennig-Thurau, Thorsten and Alexander Klee (1997), "The Impact of Customer Satisfaction and Relationship Quality on Customer Retention: A Critical Reassessment and Model Development," *Psychology & Marketing*, 14 (8), 737–64.

Hirschman, Albert O. (1970), *Exit, Voice, and Loyalty: Responses to Decline in Firms*. Cambridge, MA: Harvard University Press.

Hochschild, Arlie R. (1983), *The Managed Heart: Commercialization of Human Feeling*. Berkeley: University of California Press.

Hunt, H. Keith (1993), "CSD&CB Research Suggestions and Observations for the 1990's," *Journal of Satisfaction/Dissatisfaction and Complaining Behavior*, 6, 40–42.

Kruml, Susan M. and Deanna Geddes (2000), "Exploring the Dimensions of Emotional Labor The Heart of Hochschild's Work," *Management Communication Quarterly*, 14 (1), 8–49.

McHugo, Gregory J., Denis G. Sullivan, John T. Lanzetta, Roger D. Masters, and Basil G. Englis (1985), "Emotional Reactions to the Expressive Displays of a Political Leader," *Journal of Personality and Social Psychology*, 49 (6), 1513–29.

Morris, J. Andrew and Daniel C. Feldman (1996), "The Dimensions, Antecedents, and Consequences of Emotional Labor," *Academy of Management Review*, 21 (October), 986–1010.

Oliver, Richard L. (1981), "Measurement and Evaluation of Satisfaction Processes in Retail Settings," *Journal of Retailing*, 57 (Fall), 25–48.

——— (1997), *Satisfaction: A Behavioral Perspective on the Consumer*. New York: McGraw-Hill.

——— (1999), "Whence Consumer Loyalty?" *Journal of Marketing*, 63 (Special Issue), 33–44.

Parasuraman, A., Valarie A. Zeithaml, and Leonard L. Berry (1985), "A Conceptual Model of Service Quality and Its Implications for Future Research," *Journal of Marketing*, 49 (Fall), 41–50.

Parker, Sharon K. and Carolyn M. Axtell (2001), "Seeing Another Viewpoint: Antecedents and Outcomes of Employee Perspective Taking," *Academy of Management Journal*, 44 (December), 1085–1100.

Perdue, Barbara C. and John O. Summers (1986), "Checking the Success of Manipulations in Marketing Experiments," *Journal of Marketing Research*, 23 (November), 317–26.

Peter, J. Paul, Gilbert A. Churchill Jr., and Tom J. Brown (1993), "Caution in the Use of Difference Scores in Consumer Research," *Journal of Consumer Research*, 19 (March), 655–62.

Pugh, S. Douglas (2001), "Service with a Smile: Emotional Contagion in the Service Encounter," *Academy of Management Journal*, 44 (October), 1018–1027.

Rafaeli, Anat and Robert I. Sutton (1987), "Expression of Emotion as Part of the Work Role," *Academy of Management Review*, 12 (January), 23–37.

——— and ——— (1989), "The Expression of Emotion in Organizational Life," in *Research in Organizational Behavior*, Vol. 11, Larry L. Cummings and Barry M. Staw, eds. Greenwich, CT: JAI Press, 1–42.

Rust, Roland T. and Richard L. Oliver (2000), "Should We Delight the Customer?" *Journal of the Academy of Marketing Science*, 28 (Winter), 86–94.

——— and Anthony J. Zahorik (1993), "Customer Satisfaction, Customer Retention, and Market Share," *Journal of Retailing*, 69 (Summer), 193–215.

Salancik, Gerald R. and Jeffrey Pfeffer (1978), "A Social Information Processing Approach to Job Attitudes and Task Design," *Administrative Science Quarterly*, 23 (June), 224–53.

Schoenewolf, Gerald (1990), "Emotional Contagion: Behavior Induction in Individuals and Groups," *Modern Psychoanalysis*, 15 (1) 49–61.

Schwartz, Norbert and Gerald L. Clore (1988), "How Do I Feel About It? The Informative Function of Affective States," in *Affect, Cognition, and Social Behavior*, Klaus Fiedler and Joseph P. Forgas, eds. Toronto: Hogrefe, 44–62.

Shrout, Patrick E. and Joseph L. Fleiss (1979), "Intraclass Correlations: Uses in Assessing Rater Reliability," *Psychological Bulletin*, 86 (2), 420–28.

Stanislavski, Constantin (1965), *An Actor Prepares*. New York: Theatre Art Books.

Stone, Mervyn (1974), "Cross-Validatory Choice and Assessment of Statistical Predictions," *Journal of the Royal Statistical Society*, 36 (2), 111–47.

Surprenant, Carol F. and Michael R. Solomon (1987), "Predictability and Personalization in the Service Encounter," *Journal of Marketing*, 51 (April), 86–96.

Sutton, Robert I. and Anat Rafaeli (1988), "Untangling the Relationship Between Displayed Emotions and Organizational Sales: The Case of Convenience Stores," *Academy of Management Journal*, 31 (September), 461–87.

Taylor, Steven A. and Thomas L. Baker (1994), "An Assessment of the Relationship Between Service Quality and Customer Satisfaction in the Formation of Consumers' Purchase Intentions," *Journal of Retailing*, 70 (Summer), 163–78.

Tickle-Degnen, Linda and Robert Rosenthal (1990), "The Nature of Rapport and Its Nonverbal Correlates," *Psychological Inquiry*, 1 (4), 285–93.

Tisak, John and Carla S. Smith (1994), "Defending and Extending Difference Score Methods," *Journal of Management*, 20 (3), 675–82.

Tsai, Wei-Chi (2001), "Determinants and Consequences of Employee Displayed Positive Emotions," *Journal of Management*, 27 (4), 497–512.

——— and Yin-Mei Huang (2002), "Mechanisms Linking Employee Affective Delivery and Customer Behavioral Intentions," *Journal of Applied Psychology*, 87 (October), 1001–1008.

Verbeke, Willem (1997), "Individual Differences in Emotional Contagion of Salespersons: Its Effect on Performance and Burnout," *Psychology & Marketing*, 14 (6), 617–36.

Watson, David and Auke Tellegen (1985), "Toward a Consensual Structure of Mood," *Psychological Bulletin*, 98 (2), 219–35.

Yi, Youjae (1990), "A Critical Review of Consumer Satisfaction," in *Review of Marketing 1990*, Valarie A. Zeithaml, ed. Chicago: American Marketing Association, 68–123.

Zeithaml, Valarie A., Leonard L. Berry, and A. Parasuraman (1996), "The Behavioral Consequences of Service Quality," *Journal of Marketing*, 60 (April), 31–46.

Yong Liu

Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue

This article uses actual word-of-mouth (WOM) information to examine the dynamic patterns of WOM and how it helps explain box office revenue. The WOM data were collected from the Yahoo Movies Web site. The results show that WOM activities are the most active during a movie's prerelease and opening week and that movie audiences tend to hold relatively high expectations before release but become more critical in the opening week. More important, WOM information offers significant explanatory power for both aggregate and weekly box office revenue, especially in the early weeks after a movie opens. However, most of this explanatory power comes from the volume of WOM and not from its valence, as measured by the percentages of positive and negative messages.

Word of mouth (WOM), or "buzz," involves informal communication among consumers about products and services. Two important features distinguish WOM from other information sources, such as advertising: WOM is usually perceived as more credible and trustworthy, and it is more readily accessible through social networks (Banerjee 1992; Brown and Reingen 1987; Murray 1991). A McKinsey & Company study found that 67% of the sales of consumer goods are based on WOM (Taylor 2003).

In the arena of new product development, WOM may play particularly important roles because awareness must be built, and consumers need information when deciding whether to purchase a product they do not know well (Mahajan, Muller, and Kerin 1984). In the movie industry, which has become a popular setting to examine various new product issues (e.g., Elberse and Eliashberg 2003; Krider and Weinberg 1998; Lehmann and Weinberg 2000; Sawhney and Eliashberg 1996), hundreds of new films are released annually by Hollywood and independent filmmakers. In general, it is believed that WOM strongly influences people's movie selection (Austin 1989; Bayus 1985; Faber and O'Guinn 1984; Neelamegham and Chintagunta 1999). Indeed, the box office success of several movies, such as *The Blair Witch Project*, *My Big Fat Greek Wedding*, and *Star Wars: Episode 1–The Phantom Menace*, has been attributed to the WOM that these movies generated.

This study makes use of a data set containing actual WOM information to study the dynamic patterns of movie WOM, the explanatory power of WOM for box office revenue, and the potential antecedents of WOM. It adds to the growing literature on movie marketing and provides insights into new product development, entertainment marketing (Eliashberg and Shugan 1997), services marketing, and WOM research in general.

The WOM data were collected from the Yahoo Movies message board (see http://movies.yahoo.com/). It contains more than 12,000 posted messages, which were coded into weekly measures of volume and valence before being empirically examined. This study attempts to provide answers to the following questions: First, in markets in which the product has a low price and prerelease information is abundant (e.g., massive prerelease advertising, extensive professional reviews), how active is WOM before release, and is there any consistent shift in consumer evaluation of a new product before introduction (i.e., expectations) versus after introduction (i.e., experience based)? Second, to what extent do important measures of WOM help explain the sales of a new product, both periodically and in aggregation? Third, what factors consistently function as antecedents of WOM across time periods? If WOM is important for the success of a new product, firms should try to understand and manage these antecedents.

Advances in communications technology, especially the Internet, have made it extremely easy to share information with other people. Numerous Web sites provide message boards and information threads for users to post information and exchange opinions. This not only increases the likelihood of consumers using WOM in decision making but also provides an opportunity for researchers to gather actual WOM information. Nevertheless, few published studies have used these data, with the notable exception of Godes and Mayzlin (2004b), who study WOM of television shows from the Usenet newsgroups. Through regressions

---

Yong Liu is Assistant Professor of Marketing, Martin J. Whitman School of Management, Syracuse University (e-mail: yoliu@syr.edu). The author thanks Chuck Weinberg, Dan Putler, Tridib Mazumdar, Amiya Basu, Eunkyu Lee, and the participants at the 2004 Marketing Science Conference in Rotterdam, the Netherlands, for helpful comments on previous drafts of the article. The author also thanks Dae-Yong Ahn (doctoral student in Marketing, University of Texas at Austin) and Anindita Chakravarty (doctoral student in Marketing, Pennsylvania State University) for valuable research assistance. The author gratefully acknowledges financial support from the Snyder Innovation Management Center at Syracuse University.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

across weeks and shows, they empirically test the relationship between WOM (measured in volume and dispersion across different newsgroups) and weekly ratings. They find that the dispersion measure, but not volume, has significant explanatory power for ratings.

The current study differs from that of Godes and Mayzlin (2004b) in several ways. First, in addition to volume, the role of valence, which is coded on the basis of the content of each WOM message, is investigated in detail. Second, many of the results provide insights into the prerelease period of a new product, which is critical in markets, such as movies, in which early sales are of great concern. Third, in addition to the relationship between WOM and sales, the dynamic patterns of WOM (i.e., how it evolves and its potential antecedents) are investigated. Finally, the research context is different. Unlike network television shows, which normally run on a seasonal basis, movies have very short life cycle in theaters. They also have unique industrial features that offer interesting research questions and require special considerations in empirical analysis; for example, a lot of weight is placed on the opening week, and a minimum number of weeks are usually specified in exhibition contracts (De Vany and Walls 1999).

Theoretical Background

Although successful movies can generate great box office revenue and profits, the movie industry is also known for high risks. Of any ten major films produced, six to seven are unprofitable (Vogel 2001, p. 35; see also Shugan 1995). Two characteristics of the movie industry contribute to the popular belief that WOM influences moviegoers. First, as a prime category of popular culture goods, movies tend to receive great public interest and attention. Therefore, active interpersonal communication about movies can be expected to exist and, as indicated by the theory of information accessibility and influences (Chaffee 1982), may influence the audience. Second, the intangible and experiential nature of movie consumption makes it difficult to judge movie quality before it is actually viewed. When the alternatives are difficult to evaluate before purchase, consumers often engage in WOM to gather more information (Bristor 1990; Harrison-Walker 2001; Rogers 1983).

Volume and valence are among the most important WOM attributes that have been examined (e.g., Mahajan, Muller, and Kerin 1984; Mizerski 1982; Neelamegham and Chintagunta 1999). Volume measures the total amount of WOM interactions. Valence captures the nature of WOM messages (i.e., whether they are positive or negative). Mainly because of data availability, most extant studies examine either volume or valence, but not both. For example, Anderson (1998) and Bowman and Narayandas (2001) focus exclusively on the volume of WOM.

Two previous studies use controlled laboratory settings to investigate how WOM is related to certain behavioral features of moviegoers. In particular, Faber and O'Guinn (1984) examine how much credibility and importance college students attach to WOM versus other information sources. Burzynski and Bayer (1977) study how WOM moderates people's liking of a film. Several studies that focus on other issues of the movie market have included various proxy measures of WOM as covariates (e.g., Elberse and Eliashberg 2003; Neelamegham and Chintagunta 1999; Zufryden 1996). Two other independently conducted studies also use data from Yahoo Movies. Dellarocas, Awad, and Zhang (2004) use a Bass diffusion model to examine how user ratings posted in the opening week help explain the two Bass parameters (p = the external influence factor, and q = the internal influence factor), which are estimated from the box office history of a movie sample. They find that the volume of the first week's user ratings and their density (defined as a ratio between the volume of ratings and the first week's box office revenue), but not the numerical value of these ratings, are useful in explaining p. However, the value of user ratings becomes a significant explanatory variable for q. Duan, Gu, and Whinston (2005) use similar user-ratings data but focus on the correlation between the daily measures of these ratings and the daily box office revenue in the first two weeks. They find that user ratings have no explanatory power for box office revenue, but the volume of ratings does.

There are several important differences between the current study and those of Dellarocas, Awad, and Zhang (2004) and Duan, Gu, and Whinston (2005). First, this study disentangles the weekly effects of WOM in a way similar to Eliashberg and Shugan (1997) and Basuroy, Chatterjee, and Ravid (2003). This is useful given that different weeks are considered of different importance in the movie industry, and many managerial decisions are made on a weekly basis. Second, the data contain WOM information in the prerelease period, which can be used to explain and predict the movie's opening-week box office revenue. An important managerial value of prerelease WOM is that it can help improve the early forecast of opening-week box office and aggregate box office revenue. Third, this study investigates the potential antecedents of WOM, which provide useful implications for both WOM research and managerial practice. Finally, because the data are coded into weekly measures, it is possible to examine how WOM evolves over time.[1]

Extant studies have found that the volume of WOM correlates significantly with consumer behavior and market outcome (Anderson and Salisbury 2003; Bowman and Narayandas 2001; Van den Bulte and Lilien 2001). The reason that is often advanced to explain why the pure volume of WOM matters is consumer awareness. For example, Godes and Mayzlin (2004b) suggest that the more conversation there is about a product, the more likely someone is to be informed about it, thus leading to greater sales. On the basis of a similar rationale, research that uses diffusion models often examines WOM by either the number of adopters (Neelamegham and Chintagunta 1999) or the interaction between the number of adopters and that of nonadopters (Zufryden 1996). In this article, the effect of volume is called the "informative" effect on awareness.

---

[1] On a more technical level, the data are collected from the Yahoo Movies message board, which is characterized as a "community" site that allows bilateral discussions in a threaded manner (see http://movies.yahoo.com/). Dellarocas, Awad, and Zhang's (2004) and Duan, Gu, and Whinston's (2005) data come from the "User Review" section of Yahoo Movies, where users can post a numerical rating, a written review, or both.

Positive WOM typically gives either a direct or an indirect recommendation for product purchase. Negative WOM may involve product denigration, rumor, and private complaining. The reason valence matters is relatively straightforward; positive WOM enhances expected quality (and, thus, consumers' attitudes toward a product), whereas negative WOM reduces it. In this article, the effect of valence is called the "persuasive" effect on attitude.

Although volume and valence are among the most important measures of WOM in the literature, other measures, such as duration, intensity, and dispersion, exist (Eliashberg et al. 2000; Godes and Mayzlin 2004b). Because these measures are difficult to construct from the current data, this study focuses on volume and valence.

The effects and potential antecedents of movie WOM are investigated by examining three important aspects of a movie's run: the opening week, the weeks after opening, and aggregate box office revenue. Previous studies that have adopted similar approaches include those of Elberse and Eliashberg (2003), who separate the opening week from remaining weeks, and Eliashberg and Shugan (1997), who examine the effects of critical reviews in each of the first eight weeks after a movie's release and on aggregate box office revenue.

Comparison of WOM Before and After Release

Opposite theories exist about the degree of WOM activities before a movie's release. Several factors suggest that prerelease WOM is minimal. First, the price of watching a movie is relatively low for an average consumer; the average movie admission was $6 in 2003 (Motion Picture Association 2003). The low cost, coupled with moviegoing often being an impulse decision, implies that potential moviegoers may not go through an intensive decision-making process in which a large amount of information, such as WOM, is needed. Second, most movies are released after intensive prerelease advertising. Advertising, which is mainly done in newspapers and on television, typically amounts to 50% of the production budget (Vogel 2001, p. 96). It provides information about the opening date, main cast, some of the storyline, critical opinions (in a selected fashion), and, in many cases, whether the movie is to be released nationwide "in a theater near you" or only "in selected cities." Critical reviews also provide useful information to moviegoers and can be easily obtained from many media channels (Faber and O'Guinn 1984; Litwak 1986). Because movie critics are mostly independent from the studios, their reviews could offer supplementary value to advertising. If massive advertising and active critical reviews can meet potential moviegoers' need for information and thus be an effective substitute for WOM, the amount of WOM in the prerelease period should be minimal (Chaffee 1982). Finally, it is conceivable that the actual experience with a product may function as an important stimulus for WOM communication (Anderson 1998). Thus, the volume of WOM could be expected to be significantly large only after many moviegoers have seen the movie in the opening week.[2]

Conversely, as discussed previously, movies are popular cultural goods with an intangible/experiential nature, which increases the likelihood that prerelease WOM is active. Because WOM comes from other moviegoers, it may also be perceived as more trustworthy than advertising (Faber and O'Guinn 1984), and it may better reflect popular taste than critical reviews (Holbrook 1999). As a result, although abundant information about a to-be-released movie is provided by advertising and critical reviews, WOM may still be perceived as useful and thus serve as a complement to rather than a substitute for other information sources. This study relies on an empirical approach to examine how active WOM is during the prerelease phase of a movie and how it compares with WOM in the opening week.

By its very nature, advertising will be positive about a new movie. Notably, Eliashberg and Shugan (1997) find that critical reviews are also mostly positive. To the extent that these sources of information influence moviegoers, the valence of prerelease WOM for an average movie should be mainly positive. After a movie is released, actual experience will come into play to influence WOM valence further. Abundant research indicates that consumer satisfaction is influenced by the confirmation or disconfirmation of expectations (Anderson 1973; Oliver 1980; Olson and Dover 1979). Because the viewing public holds a comparatively high expectation for a to-be-released movie, the likelihood of disconfirmation is greater. In turn, this may result in a decrease in the positivity of WOM. Lieberman and Esgate (2002) make a similar argument for movie trailers and interactive games; that is, because trailers and games create high expectations before release, viewers are more likely to be disappointed after seeing the movie and talk unfavorably about it. Thus:

> H_1: The valence of prerelease WOM is mainly positive, and the valence of WOM in the opening week is less positive than that of prerelease WOM.

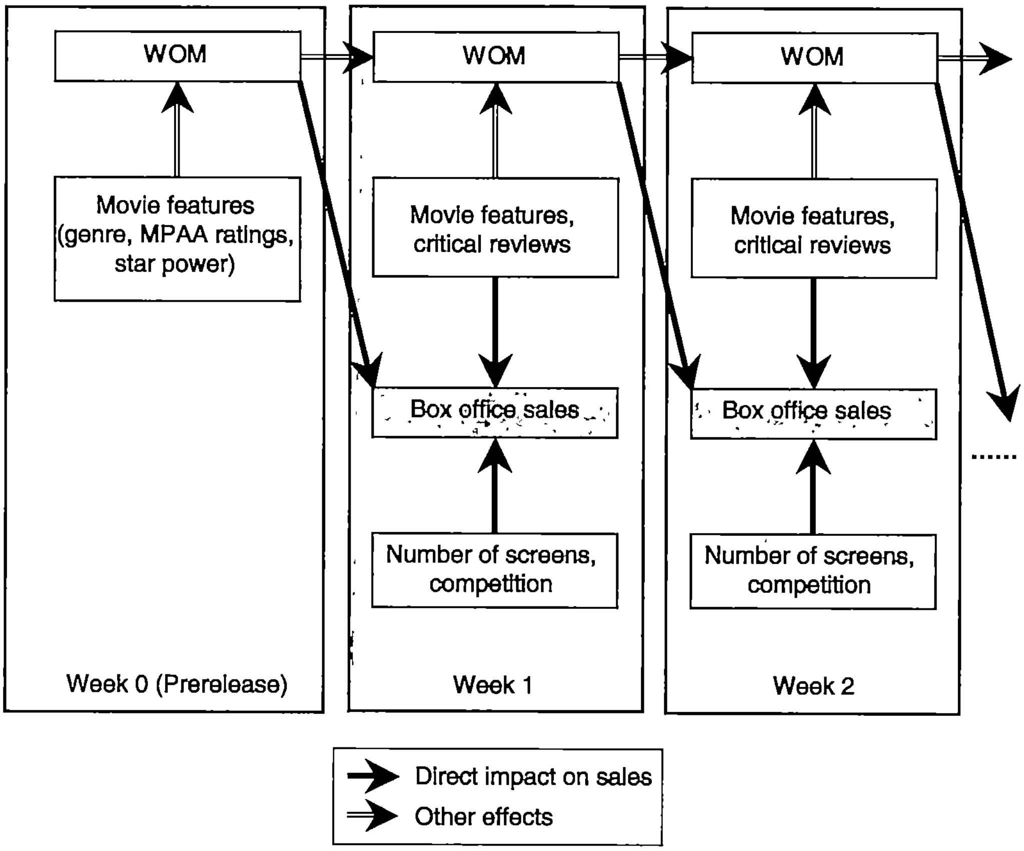
Weekly Box Office Revenues

Figure 1 presents a conceptual framework that integrates WOM into an otherwise familiar picture of the movie market (Elberse and Eliashberg 2003; Lehmann and Weinberg 2000). The extant literature consistently finds that the number of screens on which a movie is shown strongly influences box office revenue in a given week (e.g., Swami, Eliashberg, and Weinberg 1999). Evidence also indicates that critical reviews correlate with weekly box office revenue, though previous studies have reported mixed findings about the significant level of this correlation and when it exists in a movie's run (Basuroy, Chatterjee, and Ravid

---

[2] A caveat is that some audiences might have the chance to watch a sneak preview of a to-be-released movie. However, sneak previews are conducted only for selected movies, in few cities, and for a small number of people. The vast majority of moviegoers would not have experienced the movie before its release.

FIGURE 1
A Conceptual Framework of Movie Box Office Sales


WOM
Movie features (genre, MPAA ratings, star power)
Week 0 (Prerelease)
WOM
Movie features, critical reviews
Box office sales
Number of screens, competition
Week 1
WOM
Movie features, critical reviews
Box office sales
Number of screens, competition
Week 2
......
Direct impact on sales
Other effects

Notes: MPAA = Motion Picture Association of America.

2003; Eliashberg and Shugan 1997; Reinstein and Snyder 2005). Whereas the number of screens captures distribution intensity (and, thus, the audiences' accessibility to movies), critical reviews and WOM constitute most of the third-party information that moviegoers may use. Unlike critical reviews, which are static and occur mostly before the movie is released, WOM is a dynamic phenomenon that begins before release and spans the entire run. Thus, it is important to measure WOM periodically.

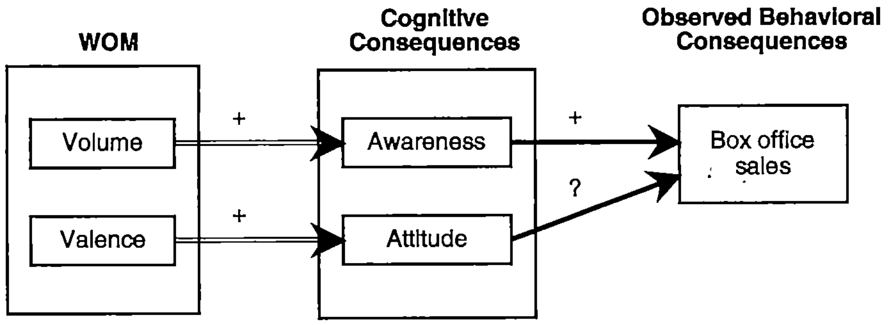
If WOM is able to influence consumer decisions, incorporating WOM into extant models that have the number of screens and critical reviews as explanatory variables (Basuroy, Chatterjee, and Ravid 2003; Eliashberg and Shugan 1997) will enhance the models' explanatory power for weekly box office revenue. This is especially true if WOM functions as a complement to critical reviews. An additional factor contributing to the explanatory power of WOM beyond that of the number of screens is the reality of overcapacity in movie theaters. The movie exhibition industry has been known for serious oversupply of theaters and seats. A series of articles that appeared in *Forbes* (Ackman 2001a, b; Schiffman 2000) estimate that the average capacity utilization of movie theaters is less than 20%. Although the number of screens picks up a large amount of variance in box office sales, the inclusion of additional factors that influence moviegoers (e.g., WOM) can explain more variance through capturing the increase in capacity utilization inside theaters.

The two measures of WOM, volume and valence, influence moviegoers through different cognition–behavior routes (Figure 2). As discussed previously, the volume of WOM mainly has an informative role to enhance consumer awareness.[3] The greater the volume of WOM, the more likely a consumer will be to hear about it. Not surprisingly, greater awareness tends to generate greater sales.

---

[3] A few studies have found that the number of arguments in a persuasive message may influence persuasion, but the nature of this influence depends on situational factors, such as the level of involvement and the strength of the arguments. For example, Petty and Cacioppo (1984) suggest that the number of arguments serves as a simple persuasive cue for low-involvement situations, but a larger number of arguments may enhance issue-relevant thinking when involvement is high. For WOM in general and for movie WOM in particular, the overall degree of consumer involvement and the factors that drive it remain to be studied. Moreover, it is an interesting issue whether the findings regarding multiple arguments in one persuasive message are applicable to a context similar to movie WOM, in which WOM messages come from different senders and usually advance different opinions. However, note that the informative effect proposed in the literature (e.g., Godes and Mayzlin 2004b; Neelamegham and Chintagunta 1999; Zufryden 1996) always occurs as moviegoers are exposed to WOM.

FIGURE 2
The Impact of WOM Measures on Potential Moviegoers



Conversely, the role of WOM valence, because of its positive/negative nature, is more of a persuasive one that influences consumer attitude. Although it is relatively straightforward to postulate that positive WOM enhances attitude and negative WOM reduces it, whether this effect can be unambiguously transformed into actual box office sales is unclear. Indeed, behavioral research has found that attitude does not always predict behavior well (Ajzen and Fishbein 1980) and that situational factors may influence behavior beyond what attitude can explain.

In the movie market, the link between attitude and behavior is further weakened when moviegoing occurs impulsively. Moreover, unlike other consumer goods that remain in the market for a long time so that consumers have sufficient time to act on the desire to purchase them, movies are in theaters only for a short time. Along this line, Sawhney and Eliashberg (1996) distinguish between the time to decide whether to watch a movie and the time to act on this decision.

Therefore, a positive relationship between WOM volume and box office revenue is postulated. However, because of the weak link between attitude and behavior, the relationship between WOM valence and box office revenue may be insignificant.

> H_2: WOM has significant explanatory power for box office revenue in the subsequent period. Whereas the effect of WOM volume should be significantly positive, that of WOM valence may be insignificant.

Aggregate Box Office Revenue

The overall performance of movies, as measured by aggregate box office revenue, is a critical issue in the industry. Beyond the weekly effects discussed in the previous section, it is useful to examine whether prerelease WOM (i.e., the buzz that accumulates before the opening week) can help explain aggregate box office revenue. The importance of accumulated buzz before a movie's release has attracted much interest in both academic research (Elberse and Eliashberg 2003) and the popular press (King 2003). If WOM is informative about weekly box office revenue and if WOM activities can correlate between adjacent weeks, prerelease WOM will help explain/predict aggregate box office revenue. In line with previous discussions, the effect of the volume of prerelease WOM is expected to be significantly positive, but that of the valence may be insignificant.

As do studies that examine aggregate sales of movies and other entertainment products (Litman 1983; Reddy, Swaminathan, and Motley 1998), this study uses factors that are constant across the movies' run as explanatory variables. In addition to prerelease WOM, this study links a movie's aggregate box office revenue to factors including star power, genre, production budget, critical reviews, and the opening strength of the movie.[4] Eliashberg and Shugan (1997) examine the effects of critical reviews on aggregate box office revenue and find that though the percentage of positive reviews is a significant predictor, the volume of reviews is not. The role of star power, genre, and production budget has been examined in a few studies, but the results are mixed (Basuroy, Chatterjee, and Ravid 2003; Elberse and Eliashberg 2003; Litman 1983).

Opening strength refers to the number of screens in the first few weeks of a movie's run. A wide opening strategy may generate higher total sales because greater accessibility better appropriates the public interest in a movie that has accumulated before release. However, opening wide may also merely shift future demand into early weeks if the overall demand is not increased. If audiences are not stimulated enough by a wide opening to go to theaters early and if the number of screens drops quickly, total sales will be negatively influenced. The business press suggests that movies are increasingly opening wide (King 2001).

> H_3: Prerelease WOM offers significant explanatory power for aggregate box office revenue. However, most of the explanatory power comes from its volume, not its valence.

---

[4] Other variables, such as Motion Picture Association of America (MPAA) ratings, directors, distributors, and whether the movie is a sequel, are potential explanatory variables for aggregate box office revenue. However, most studies failed to find these variables to be significant (e.g., Elberse and Eliashberg 2003; Litman 1983). These were also included in the estimation and were found to be insignificant as well. Thus, they are not included in the analyses reported here.

Potential Antecedents of WOM

The double-lined arrows in Figure 1 illustrate the factors that are potential antecedents of WOM. These include five groups of factors: movie genres, Motion Picture Association of America (MPAA) ratings, whether major stars are in the cast, critical reviews, and the WOM measures in the previous week. Movie genres (e.g., action, drama, horror) classify movies on the basis of their storyline, and MPAA ratings (G, PG, PG-13, and R) classify movies on the basis of content, primarily according to the suitability of the movie to different audience segments. As a communication process, WOM activities can be influenced by the strength, inherent interestingness, and possibly the controversy of the topic. Movies with certain genres (e.g., action, adventure) and MPAA ratings (e.g., R, which requires a parental guardian for moviegoers under the age of 17) may be associated with distinctive patterns of WOM behavior, especially during the prerelease period when the curiosity about movies is high.

The phenomenon of star power is substantial in the entertainment sector. Movie stars are the frequent topic of interpersonal communication among moviegoers. Thus, the existence of stars is likely to correlate with WOM activities. Because the literature is inconclusive about the role of movie stars on box office revenue and because Hollywood seems to promote a culture of stars despite the high pressure faced by studios to control production budgets (Vogel 2001), either a confirmation or a disconfirmation of the impact of star power on WOM would be valuable.

Critical reviews are usually published in major newspapers, in magazines, and, increasingly, on movie Web sites. They may influence moviegoers' expectations of a film and thus correlate with WOM. However, such influence may exist only if the majority of movie audiences read these reviews. Because there is no published evidence about the size or the demographic features of those who read critical reviews, it is difficult to speculate on the significance of the correlation between reviews and WOM.[5] Although most critical reviews begin circulating before a movie's opening week, the data do not contain information about the exact time the reviews became available. To reduce any ambiguity in the analysis for the prerelease period, critical reviews are included as a potential antecedent of WOM only in a movie's opening week and the weeks after it.

Extant studies advance the theory that WOM can be viewed as a consequence of social networks; thus, it is contagious and persistent (Banerjee 1992; Brown and Reingen 1987; Murray 1991). This indicates that on the aggregate level, WOM in adjacent weeks will be positively correlated. From a behavioral point of view, active WOM during a period helps promote awareness and interest in a movie; therefore, more people will see the movie in the next period and subsequently talk about it. This can happen for two reasons: First, greater WOM makes the memory of the movie last longer, thus increasing the likelihood that a moviegoer will talk about the movie after seeing it. Second, the perceived social benefit of talking about a popular subject (i.e., movie) tends to be greater than talking about a less popular one (Brown and Reingen 1987).

The data can be used to examine the aggregate effect of this temporal correlation. That is, a week with greater WOM activities is more likely to be followed by another week of active WOM than is a week with less WOM. This phenomenon is called a "carryover" effect of WOM, which, in the context of this study, refers to the possibility that WOM activities in a week correlate with WOM in the subsequent week.[6]

> H_4: Movie genres, MPAA ratings, star power, and critical reviews are among the potential antecedents of movie WOM. A dynamic carryover effect exists between WOM activities in adjacent weeks.

Empirical Analysis and Discussion of Results

Data and Variables

The WOM data are collected from the Yahoo Movies message board (http://movies.yahoo.com/). There are several reasons Yahoo Movies serves as a good source of movie WOM. First, it is one of the most popular movie Web sites. *Variety* (Graser 2002) reports that Yahoo Movies has the largest number of unique visitors among its Web competitors. Second, Yahoo Movies requires no access fee for either browsing or posting a message. This helps reduce any possible bias in the demographic composition of the Web site's visitors. Third, the structure of the Web site is well designed so that finding and collecting information is straightforward, thus reducing possible errors during data collection. Finally, WOM messages are archived and indexed numerically by the dates when they were posted. Thus, it is possible to track the period in the movie's run to which a particular message belongs.[7] The prerelease period is from the date when the first WOM message about a movie is posted to the date before the official movie release. It captures the accumulation of WOM before the movie opens.

---

[5] Movie advertisements also display critical reviews from time to time. However, such displays are highly selective; only positive evaluations, such as "two thumbs up," are displayed. To the extent that consumers understand this selective process, the impact of critical reviews through movie advertisements should decrease (Taylor 2003).

[6] Note that this carryover effect does not require the volume of WOM to keep increasing over time. There are other factors that may function as antecedents of WOM, such as weekly advertising spending and whether the movie is a sequel. The data do not contain sufficient information or variance to investigate these factors. This is an issue for further research.

[7] A few other Web sites (e.g., IMDB.com) also provide user comments about movies. However, these Web sites do not attract as many visitors as Yahoo Movies (Graser 2002). Furthermore, in general, they do not provide functions similar to Yahoo Movies, which allows threaded message posting and replying. Collecting and organizing the postings from these other sites are more difficult. For example, the user comments on IMDB.com are not numbered, thus increasing the difficulty of an already tedious data-processing task. To the extent that no significant difference exists

Coding WOM messages into valence is an extremely tedious task that poses serious constraint on the amount of data that can be studied. To cope with this difficulty, this study focuses on movies released during a five-month period (May–September) in 2002. After excluding titles for which there is no reliable information about all control variables, the final data set contains 40 movies, with 12,136 WOM messages to be coded.

Three judges independently read each of the 12,136 messages and assigned them to one of five categories: positive, negative, mixed, neutral, and irrelevant. The messages classified as positive or negative either show clear overall assessment of the movie or provide direct recommendations. A message is classified as mixed if it expresses positive opinions about some aspects of the movie but negative opinions about others and if there is no clear overall assessment. A message is classified as neutral if it talks about the movie but does not provide any positive or negative comments. Finally, a message is classified as irrelevant if it is not related to the movie itself. The three independent codings are integrated using the majority rule: If at least two judges assign the same category, that category is used for the message. If all three judges disagree, the message is coded as disagree. The level of disagreement among the three judges was very low (approximately 4% for all the weeks examined). For example, for WOM posted during the opening week, the judges did not achieve a majority agreement for only 4.1% of all messages. The data require the use of aggregate measures for WOM valence. Consistent with previous studies that involve valence coding (Basuroy, Chatterjee, and Ravid 2003; Eliashberg and Shugan 1997), the valence of WOM is measured in two dimensions: the percentage of positive messages and the percentage of negative messages.[8]

Other data come from various public sources. The weekly box office revenue and number of screens are collected from *Variety* magazine, as are the critical reviews. Eliashberg and Shugan's (1997) method to calculate the total number of reviews, the percentage of positive reviews (CRPRO), and the percentage of negative reviews (CRCON) is used. Similar to their data, an extremely high correlation exists between CRPRO and CRCON (–.878, $p <$ .0001). As a result, these data must be included separately in the estimation. Production budget comes from another popular movie Web site, The Numbers (see www.the-numbers.com), and MPAA ratings and movie genres are collected from the Internet Movie Database (http://us.imdb.com). Star power is based on the star list published annually by *Premier*; a movie is classified as having star power if at least one of its cast members appears in "The Power List 2002." Table 1 provides the basic summary sta-

TABLE 1
Key Summary Statistics of the Movie Sample

Variable	Median	M	SD
Box office (aggregate)	$41,986,198	$67,005,348	$56,681,186
Ratio of opening week to aggregate box office	43.1%	41.0%	14.3%
Exhibition longevity (in weeks)	15.0	15.5	4.89
Production budget	$48,000,000	$51,650,000	$30,955,344
Screens in opening week	2633.5	2518.9	890.3
Volume of total WOM messages	139.5	303.4	411.8
Volume of WOM (prerelease)	45.5	84.6	89.1
Percentage of positive WOM (prerelease)	27.3%	28.5%	12.5%
Percentage of negative WOM (prerelease)	9.1%	10.5%	8.2%
Volume of WOM (opening week)	49.0	94.8	134.5
Percentage of positive WOM (opening week)	20.3%	23.8%	11.9%
Percentage of negative WOM (opening week)	13.3%	16.1%	13.5%
Volume of critical reviews	16.5	16.3	4.1
Percentage of positive reviews	28.1%	36.9%	28.7%
Percentage of negative reviews	32.5%	37.6%	30.6%

Notes: Total number of movies = 40. Total number of WOM messages analyzed = 12,136.

---

in the demographic profiles of visitors to these movie Web sites, using data from Yahoo Movies can be fairly representative. Nevertheless, it is acknowledged that compiling data from different Web sites and examining their differences and the collective influence on product sales could be potentially useful.

[8] An alternative method is to use the absolute numbers of positive and negative messages, either individually or together with the volume of messages. However, note that doing so does not clearly separate the volume effect from the valence effect and may produce spurious results, because both numbers are significantly correlated with the total volume. For example, using the number of negative reviews to explain box office revenue always yields a positive sign, indicating that the effect of total volume may dominate the valence effect. Using the numbers of positive and negative messages together with the volume produces an unacceptable variance inflation factor, a strong signal for multicollinearity. Therefore, in this study, the analysis and discussion are restricted to the volume measure and the valence measures that are defined as percentages.

tistics of the sample.[9] Following the work of Eliashberg and Shugan (1997) and Basuroy, Chatterjee, and Ravid (2003), and in light of the varying exhibition longevities among the 40 movies, empirical analyses are conducted for the first eight weeks in the movies' run. This covers most of the interesting periods for the movies; the first eight weeks account for 97% of the box office revenue in the data.

Dynamic Features of Movie WOM

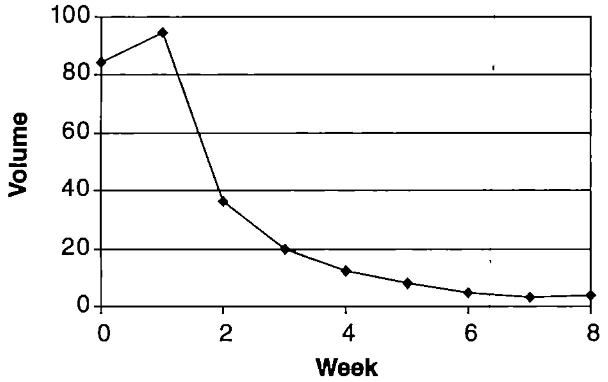
For illustrative purposes, Figure 3 plots how volume and valence change over time, beginning with the prerelease period (Week 0), then the opening week (Week 1), and so on. Figure 4 further provides plots for two specific movies, *xXx* and *Scooby-Doo*, to illustrate different patterns of the valence measures.

Figure 3, Panel A, reveals that movie WOM is the most active during the prerelease period and the opening week, and then gradually decreases. For the 40 movies in the sample, most movies (i.e., 23) have their greatest WOM volume during the prerelease period, and 15 movies have the most active WOM during the opening week. Compared with all other periods, the opening week has the greatest volume of WOM; on average, 94.8 WOM messages per movie are posted. The volume of prerelease WOM is slightly lower but is still very high; on average, 84.6 messages are posted during this period. The average volume drops significantly to 36.6 in the second week and continues to decrease.

Therefore, although other information sources are abundant and most moviegoers have not watched the movies, WOM is very active in the prerelease period. It suggests that WOM is more of a complement to other information sources than a substitute. This finding lends direct support to the hypothesis that WOM can be a useful measure to explain box office sales. It also implies that instead of being

FIGURE 3
Dynamic Patterns of WOM for the Movie Sample

A: The Average Volume of WOM



B: The Average Valence Measures

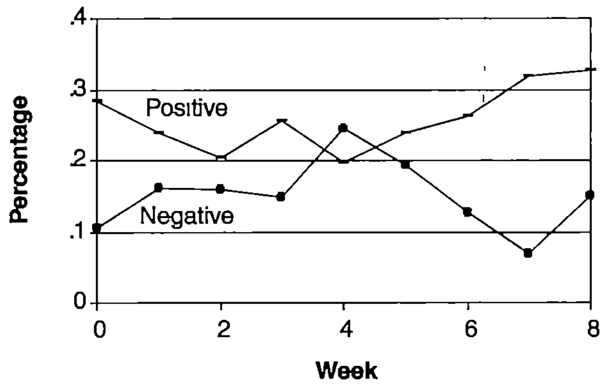
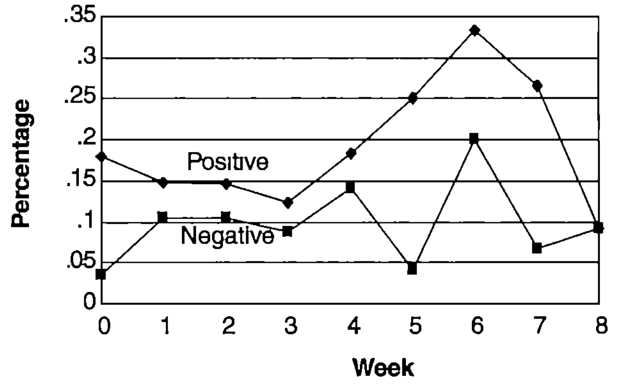
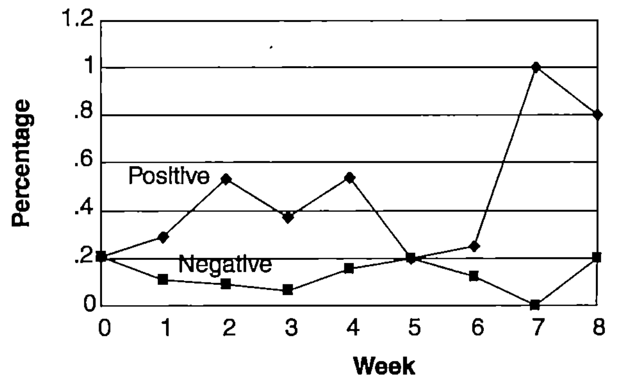


FIGURE 4
Valence of WOM for Movies *xXx* and *Scooby-Doo*

xXx



Scooby-Doo



---

[9] Note that for both WOM and critical reviews, the percentage of positive comments and the percentage of negative comments do not add up to 100%, because other categories (e.g., mixed) exist when the valence measures are constructed. It is also interesting that WOM is less often positive or negative than critical reviews. This happens because critical reviews tend to give clear overall assessments, but WOM is more likely to be ambiguous.

based on actual experiences with a movie, prerelease WOM can be fairly speculative in nature.

Figure 3, Panel B, indicates that the valence measures have a different pattern from that of volume. Although there are variations in positive and negative percentages from week to week, such changes in terms of direction and magnitude appear to be rather random. Furthermore, the average percentage of positive WOM is greater than that of negative WOM in the early periods, especially during prerelease. To quantify this finding further, each movie is examined individually. During the prerelease period, the percentage of positive WOM is greater than the percentage of negative WOM by a margin of 5% for as many as 31 movies. These margins are within 5% of each other for 8 movies, and only in one case (for the movie *The Banger Sisters*) does the percentage of negative WOM exceed that of positive WOM by 5%.

Figure 3, Panel B, also indicates that movie audiences' expectations about a typical movie (before release) are usually higher than what they think in the opening week. In particular, the percentage of positive WOM begins at a comparatively high level during prerelease and drops to a lower level in the opening week. At the same time, the percentage of negative WOM begins at a comparatively low level during prerelease but increases to a higher level in the opening week. Among the movies in the sample, the WOM for about half (19 movies) drops by more than 5% from prerelease to the opening week. The other movies either have comparable levels of positive WOM between the two periods (11 movies) or have more positive WOM during the opening week (10 movies), and 17 movies have their negative WOM increase by more than 5% from prerelease to the opening week (e.g., *xXx*), whereas only 6 movies have noticeable decreases in negative WOM (e.g., *Scooby-Doo*).

Thus, there is support for H_1, which postulates that WOM tends to be positive in the prerelease period but becomes less so in the opening week. In addition to illustrating the confirmation or disconfirmation theory of expectations in the context of movies, these results imply a potential drawback of intensive advertising. That is, advertising may generate awareness for new products, but it may also attract nontarget consumers who are more likely than others to spread negative WOM. In the movie market, these are the people who normally would not like a particular movie. This problem is of greater concern when retailers (theaters) would have kept the product longer if WOM were more positive. As a result, seeking ways to make advertising more targeted should be of considerable value. Without the concern about negative WOM, a firm would want the "wrong" buyers to buy because they still generate additional revenue.

The remaining part of this section discusses three sets of analyses of how the measures of WOM help explain weekly and aggregate box office revenue and what constitutes potential antecedents of WOM. To cope with potential nonlinearity, double log specifications are used for the regression equations; natural logarithms are taken for variables that could take unbounded positive values. The terms $LNMSG_t$, $POSIPER_t$, and $NEGPER_t$ are used to denote volume (in natural logarithms) and valence (in percentages) of WOM in week t (t = 0 indicates the prerelease period). Two weekly variables are included to control for the effects of competition; one is the number of new releases among the top 20 movies, and the other is the average age, measured in weeks, of the top 20 movies (for a discussion about why these factors may capture the effect of competition, see Elberse and Eliashberg 2003). In all estimations, multicollinearity is tested for by calculating the variance inflation factor. The variables used produced acceptable variance inflation factor values. Table 2 lists all the variables, and Table 3 provides the correlation matrix among WOM measures for a movie's prerelease and opening week. Similar

TABLE 2
List of Variables Used in the Analyses

Variable Name	Meaning
t (subscript)	Week number in movie's life span (t = 0 is prerelease)
CONST	The constant intercept in the regression models
CRCON	Percentage of negative critical reviews
CRPRO	Percentage of positive critical reviews
GACADV	Movie genre is action/adventure
GCOMEDY	Movie genre is comedy
GSCIFI	Movie genre is science fiction
$LNAGE_t$	Average age of the top 20 movies in week t (in natural log)
LNCRITIC	Number of critical reviews (in natural log)
$LNMSG_t$	Number of WOM messages in week t (in natural log)
$LNNEW_t$	Number of new releases among the top 20 movies in week t (in natural log)
LNPROD	Value of production budget of the movie (in natural log)
$LNREV_t$	Box office revenue in week t (in natural log)
$LNSCRN_{1-4}$	Number of screens in the first four weeks after opening (in natural log)
$LNSCRN_t$	Number of screens in week t (in natural log)
MPAA-PG	Movie is rated PG in MPAA ratings
MPAA-R	Movie is rated R in MPAA ratings
$NEGPER_t$	Percentage of negative WOM in week t
$POSIPER_t$	Percentage of positive WOM in week t
STAR	There are major stars in the movie

TABLE 3
Correlation Matrix for Prerelease and Opening Week WOM

	$LNMSG_0$	$POSIPER_0$	$NEGPER_0$	$LNMSG_1$	$POSIPER_1$	$NEGPER_1$
$LNMSG_0$	1.000	—	—	—	—	—
$POSIPER_0$	–.312	1.000	—	—	—	—
	(.050)					
$NEGPER_0$	.026	–.265	1.000	—	—	—
	(.873)	(.098)				
$LNMSG_1$	.764	–.034	–.007	1.000	—	—
	(.000)	(.835)	(.965)			
$POSIPER_1$	.053	.211	–.197	–.025	1.000	—
	(.746)	(.191)	(.223)	(.880)		
$NEGPER_1$	–.216	.036	.267	–.071	–.426	1.000
	(.180)	(.825)	(.096)	(.664)	(.006)	

Notes: *p* values are in parentheses.

patterns exist for the remaining weeks. The most significant correlation exists between $LNMSG_1$ and $LNMSG_0$, implying the possibility of a carryover effect.

WOM and Weekly Box Office Revenue

To examine the weekly effects of WOM, the following model is estimated:

$$(1)\quad LNREV_{it} = CONST + a_1LNSCRN_{it} + a_2LNMSG_{i,t-1} + a_3POSIPER_{i,t-1} + a_4NEGPER_{i,t-1} + a_5LNCRITIC_t + a_6CRPRO_t + a_7LNNEW_t + a_8LNAGE_t + \varepsilon_{it},$$

for each week separately (t = 1, 2, ..., 8), where i indexes the movies (i = 1, 2, ..., 40) and is omitted hereafter to avoid notation clutter. The covariates include the number of screens ($LNSCRN_t$), critical reviews (LNCRITIC, CRPRO), and the degree of competition ($LNNEW_t$, $LNAGE_t$). This specification is similar to that of Eliashberg and Shugan (1997) and Basuroy, Chatterjee, and Ravid (2003), except for the additional (lagged) WOM measures and the two competition variables.[10]

Table 4 reports the weekly regression results. All the regressions achieved high levels of model fit, and all the significant parameters have expected signs. For the opening week, the joint F test shows that the measures of WOM have significant explanatory power for box office revenue. However, whereas the volume of WOM ($LNMSG_0$) is highly significant ($p < .001$), the valence measures ($POSIPER_0$ and $NEGPER_0$) are not. In other words, almost all the explanatory power that WOM brings to the estimation comes from its volume but not its valence. The estimation for the next four weeks (Week 2–Week 5) yields the same results; that is, the volume of WOM ($LNMSG_{t-1}$) is highly significant, but the valence ($POSIPER_{t-1}$ and $NEGPER_{t-1}$) is not. At Week 6, the effect of volume becomes marginal and then disappears. This happens because WOM activities are gradually reduced as people lose interest in a movie during subsequent weeks. Thus, the increasingly sparse and irregular WOM fails to provide informative measures for weekly box office revenue. For example, there are only 5.1 WOM messages posted per movie in Week 6, compared with 94.8 in the opening week. As a result, H_2 is supported, but only for the early weeks. This is particularly enlightening because most of the box office sales occur in these weeks; for the movies in the data, the first five weeks contribute to 91.3% of aggregate box office revenue.[11]

To examine the robustness of these results, especially the insignificance of valence, which might be particularly surprising, several further analyses were conducted. A typical reason for the lack of significance is that the sample size may not have been sufficiently large. Whereas the number of movies that can be analyzed is practically constrained by the labor-intensive coding process, the sample size is still comparable to those of extant studies. For example, Godes and Mayzlin (2004b) analyze 41 television shows; Lehmann and Weinberg (2000) use 35 movie and video titles; Eliashberg and Shugan (1997) analyze 56 movies; and Dellarocas, Awad, and Zhang (2004) use 25 movies in a ratings survey and then 80 movies for box office analysis (40 movies for estimation and 40 for prediction). Nevertheless, a pooled regression was conducted to examine whether

---

[10] In an effort to test and control for any effects that might be caused by unobservable variables, the estimation of a set of movie-specific information, such as movie genres, MPAA ratings, production budget, whether a movie is a sequel, and whether there are major stars or directors, is included. However, these variables are mostly insignificant. Thus, they are omitted from the final model of estimation.

[11] A question here is whether the volume of WOM is merely picking up the effects of advertising. To examine this issue, data on the advertising budget for each movie were collected and included in weekly regressions with WOM. The estimation results remain similar, and the volume of WOM is still highly significant. Thus, movie WOM offers explanatory power for weekly box office revenue beyond that of advertising.

TABLE 4
The Effects of WOM Communication on Weekly Box Office Revenue

Week (t)	1	2	3	4	5	6	7	8
CONST	7.273	9.346	6.449	8.171	6.469	4.981	.594	4.067
	(.000)**	(.000)**	(.000)**	(.000)**	(.000)**	(.015)**	(.000)**	(.317)
$LNSCRN_t$	.566	.401	.673	.985	.973	1.050	1.379	1.185
	(.000)**	(.006)**	(.000)**	(.000)**	(.000)**	(.000)**	(.000)**	(.000)**
$LNMSG_{t-1}$	.592	.345	.366	.275	.387	.169	–.088	.144
	(.000)**	(.000)**	(.000)**	(.005)**	(.000)**	(.0103)	(.449)	(.362)
$POSIPER_{t-1}$	.784	–.300	.812	.186	.342	.045	–.185	.118
	(.257)	(.674)	(.133)	(.603)	(.312)	(.910)	(.497)	(.715)
$NEGPER_{t-1}$	.059	–1.369	–.610	.445	.385	–.173	–.265	–.605
	(.952)	(.092)*	(.264)	(.309)	(.136)	(.620)	(.399)	(.655)
LNCRITIC	.759	1.106	.894	.343	.421	.657	.358	.561
	(.018)**	(.001)**	(.020)**	(.323)	(.137)	(.192)	(.228)	(.369)
CRPRO	.695	.398	.688	.728	.736	.340	.386	.307
	(.019)**	(.235)	(.076)*	(.047)**	(.009)**	(.336)	(.288)	(.454)
LNNEW	–.111	.005	.153	–.251	–.136	–.086	.073	–.112
	(.505)	(.984)	(.494)	(.127)	(.423)	(.738)	(.711)	(.682)
LNAGE	.146	–.339	.022	–1.378	–.681	–.039	–1.510	.060
	(.760)	(.516)	(.972)	(.045)**	(.140)	(.964)	(.097)*	(.963)
Model fit F	44.87	16.86	22.70	27.50	50.37	23.58	69.04	22.45
(*p* value)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
Adjusted R^2	.902	.765	.820	.855	.925	.862	.954	.905
Joint F test of WOM measures	14.01 (.000)	9.64 (.000)	10.87 (.000)	4.01 (.017)	10.48 (.000)	1.07 (.382)	.43 (.735)	.53 (.672)

$^*p < .10$
$^{**}p < .05$.
Notes: The dependent variable is $LNREV_t$. *p* values are in parentheses.

the insignificance of WOM valence can be attributed to sample size. To do so, all the weekly data of the 40 movies were pooled to regress weekly box office revenue on the explanatory variables. With 378 observations, the pooled regression achieved a high level of model fit (adjusted R^2 = .95). However, the results remain the same; that is, the volume of WOM is highly significant, but valence is not.[12]

It is also possible that the valence of WOM works differently for movies that receive mostly positive WOM and those that receive mostly negative WOM. A relevant argument is the diminishing impact of incremental WOM messages. If the overall WOM information about a movie is positive, the addition of another positive message may not be as influential as the addition of a negative one. The reverse is true for a movie that has overall negative WOM. Thus, estimating all movies together might have neutralized the impact of the valence measures. This argument motivated an examination of the impact of WOM by conducting separate regressions on different movie groups (especially those that received mostly positive WOM and those that received mostly negative WOM). The results of WOM volume and valence remain unchanged.

In addition to the examination of weekly box office revenue, an analysis was conducted for aggregate box office revenue using prerelease WOM and other control variables; valence measures were again found to be insignificant. Finally, various model specifications, such as linear and semilog, were examined, and all lead to qualitatively similar results.

Therefore, the results confirming H_2 appear to be fairly robust. They suggest that though WOM is a useful information source in the movie market, it correlates with box office revenue mainly through an informative role on awareness but not a persuasive path of attitude changes. In addition to the theoretical reasons discussed previously regarding the weak correlation between attitude and behavior, whether a person likes a particular movie is (to a large extent) an issue of personal taste, which lends additional support to this finding. It is clear that WOM increases awareness among potential moviegoers, which is confirmed by the highly significant effect of volume, but when it comes to deciding whether to watch a particular movie, to what extent should another person's subjective assessment be reflective of the WOM recipient's own taste? If the belief that a bad (good) movie for one person may be enjoyable (not enjoyable) to another is sufficiently strong, or if the WOM receiver attributes the valence of WOM to the taste of the message sender (see Eagly, Wood, and Chaiken

---

[12] Note that a pooled regression, though it offers a greater number of observations and thus increases the chance of a parameter to be significant, has its limitations. It estimates the average effect across all periods. Thus, it is not particularly useful if there is a significant difference of the estimated effects in different periods.

TABLE 5
Prerelease WOM and the Aggregate Box Office Revenue (LNGROSS)

Independent Variables	Parameter Estimates	t-Value (p Value)
CONST	8.875	3.95 (.000)**
$LNSCRN_{1-4}$	516	2.50 (.018)**
LNPROD	–.015	–.10 (.921)
$LNMSG_0$	.589	4.65 (.000)**
$POSIPER_0$	.342	.44 (.667)
$NEGPER_0$	–.515	–.44 (.665)
LNCRITIC	.579	1.74 (.093)*
CRPRO	1.069	3.21 (.003)**
STAR	.168	.96 (.346)
GSCIFI	–.194	–.84 (.409)
GACADV	–.164	–.86 (.394)

Joint test ($LNMSG_0 = POSIPER_0 = NEGPER_0 = 0$): F = 9.49 ($p < .001$)
Model fit: F = 13.79 ($p < .001$), adjusted $R^2 = .77$

*$p < .10$.
**$p < .05$.

[1978] and other studies on attribution theory), the valence will lose much of its influence. This explanation is consistent with several previous studies on WOM. For example, Russell, DeCarlo, and Ramaswami (2001) find that when message receivers attribute the negativity of WOM to the communicator, brand evaluations actually increase.

The finding that WOM operates mainly through increased awareness and the argument that the experiential/intangible nature of movie consumption enhances the need for WOM lead to an interesting prediction; namely, the effect of WOM on to-be-released sequel movies should be less than that on nonsequels because moviegoers have more information about and experience with sequels as a result of the mostly successful prequels. This difference should be the most obvious in the opening week when box office revenue is the greatest and when the WOM effect is highly significant.

To examine this prediction, regressions similar to those reported in Table 4 could be conducted for sequel and nonsequels separately. However, the small number of sequels in the data (there were only 3) prevents estimations for completely separated samples. Thus, a different test was used. In particular, the opening-week regression was conducted for nonsequels only (37 movies), and whether the effect of volume was greater than that based on all 40 movies was observed. The parameter $LNMSG_{t-1}$ was .6044 for nonsequels but .5917 for all movies. Because the estimations are based on double log specifications, this seemingly small difference actually implies that the average volume of WOM during prerelease yields 6% more box office revenue in the opening week for nonsequels than for all movies combined (i.e., $84.6^{6044 - 5917} = 1.06$). For the average opening-week revenue of $27.01 million in the sample, this is a $1.62 million difference. This difference is especially significant considering that the test is less powerful than a direct comparison between sequels and nonsequels and that it is based on merely 3 of the 40 movies.

For the early weeks, inclusion of the WOM measures produced noticeably better model fit than the previous studies that used only the remaining explanatory variables. From Week 1 to Week 5, the average enhancement of model fit measured by adjusted R^2 is 17%.

WOM and Aggregate Box Office Revenue

Table 5 presents the results of regressing aggregate box office revenue (LNGROSS) on prerelease WOM and other covariates. The opening strength of movies is measured as the total number of screens in the first four weeks; this accommodates the industry practice of adopting a minimum number of screens in exhibition contracts for the first few weeks (for similar issues, see De Vany and Walls 1999) and reduces any possible biases associated with the movies that do not achieve the widest distribution in the opening week. The $LNSCRN_{1-4}$ turns out to be highly significant; the wider distribution a movie has in the first four weeks, the greater the aggregate box office revenue will be. The percentage of positive reviews (CRPRO) has a significant, positive impact. The volume of critical reviews (LNCRITIC) is also significant, but only marginally. These results are similar to those that Eliashberg and Shugan (1997) report. Movie stars (STAR) or movie genres are not found to be informative of aggregate box office revenue.

The three measures of WOM are highly significant as a group, confirming the value of using prerelease buzz to explain overall box office performance. As with the weekly analysis, whereas the volume of WOM ($LNMSG_0$) is the most significant among all explanatory variables, the valence measures ($POSIPER_0$ and $NEGPER_0$) are not significant. Thus, H_3 is strongly supported. An important managerial value of this finding is that the amount of prerelease buzz can be a useful indicator of the ultimate box office performance. The effectiveness of box office forecast models (e.g., Sawhney and Eliashberg 1996) can be improved by including actual prerelease WOM information.

Potential Antecedents of WOM

As discussed previously, prerelease WOM can be related to three groups of factors: movie genres, MPAA ratings, and whether major stars are in the cast. Beginning with the opening week, WOM can also be related to critical reviews and the WOM in the previous week. Because volume is the most critical factor that correlates with box office revenue, it is the focus of the subsequent analyses. Analyses for the valence measures ($POSIPER_t$ and $NEGPER_t$) were also conducted, but they could not be consistently explained by various models in almost all the periods. Table 6 summarizes the results.

All the regressions for prerelease and the eight weeks achieved significant model fit. During the prerelease period, it appears that action and adventure movies tend to receive higher-than-average WOM communication, whereas R-rated movies tend to receive less. Beginning with the opening week, the only explanatory variable that is consistently significant is the volume of WOM in the previous week ($LNMSG_{t-1}$). This shows the carryover effect such that a

TABLE 6
Potential Antecedents of the Volume of WOM

Week (t)	0 (Pre-release)	1 (Opening Week)	2	3	4	5	6	7	8
CONST	3.817	.615	–.243	1.380	–.456	–4.320	4.086	4.351	.066
	(.000)**	(.680)	(.868)	(.323)	(.838)	(.164)	(.133)	(.326)	(.911)
GACADV	555	–.143	–.075	–.448	–.152	.295	–.181	.340	—
	(.045)**	(.615)	(.764)	(.046)**	(.688)	(.444)	(.549)	(.467)	
GCOMEDY	–.330	–.358	.343	–.633	.458	.136	.140	.601	—
	(.279)	(.278)	(.231)	(.008)**	(.284)	(.712)	(.734)	(.496)	
MPAAPG	.668	–.684	.343	.418	–.303	.618	.095	–1.823	–.834
	(.104)	(.141)	(.364)	(.192)	(.597)	(.258)	(.892)	(.237)	(.197)
MPAAR	–.913	–.322	.753	–.346	.408	–.047	–.098	.225	–.452
	(.009)**	(.370)	(.028)**	(.165)	(.377)	(.913)	(.811)	(.720)	(.463)
STAR	.408	.343	–.109	.469	.136	–.145	.806	.558	—
	(.162)	(.269)	(.694)	(.068)*	(.765)	(.762)	(.095)*	(.489)	
LNMSG$_{t-1}$	—	1.169	.929	.715	.986	1.037	.625	.915	.595
		(.000)**	(.000)**	(.000)**	(.000)**	(.000)**	(.000)**	(.001)**	(.031)**
POSIPER$_{t-1}$	—	1.997	–1.148	1.149	.102	1.984	–1.888	1.206	2.067
		(.125)	(.309)	(.114)	(.917)	(.048)**	(.014)**	(.340)	(.049)**
NEGPER$_{t-1}$	—	2.405	–2.085	–1.988	–.735	1.889	–.698	2.603	.512
		(.190)	(.098)*	(.008)**	(.602)	(.041)**	(.236)	(.349)	(.781)
LNCRITIC	—	–.834	–.115	–.442	.039	.808	–1.375	–2.211	—
		(.132)	(.810)	(.367)	(.963)	(.476)	(.203)	(.275)	
CRPRO	—	.792	.302	.897	–.965	.929	1.099	1.759	—
		(.165)	(.573)	(.038)**	(.214)	(.182)	(.220)	(.373)	
Model fit F	4.36	7.53	9.78	16.22	3.69	5.78	5.75	3.41	2.89
(p value)	(.004)	(.000)	(.000)	(.000)	(.006)	(.001)	(.003)	(.094)	(.067)
Adjusted R^2	.301	.626	.698	.813	.464	.648	.683	.617	.371

*$p < .10$.
**$p < .05$
Notes: The dependent variable is LNMSG$_t$. p values are in parentheses. In Week 8, movie genre, star power, and critical reviews are excluded because of high multicollinearity.

week of active WOM is likely to be followed by another week of active WOM. However, consistent patterns linking WOM valence in a previous week to WOM volume in the current week were not found. Table 6 also indicates that star power is insignificant in explaining the volume of WOM during the prerelease period and most of the subsequent weeks. In addition, WOM is not significantly influenced by critical reviews, confirming the previous result that WOM and critical reviews are more of complements to each other than substitutes.[13]

Regarding H$_4$, consistent support was found for the carryover effect. The effects of other factors (e.g., movie genre, MPAA ratings) are mixed and mostly insignificant. Managerially speaking, these results point to the challenge that studios may face when trying to create and manipulate WOM by working on factors such as star power and movie genres (Rosen 2000). They also point to the importance of managing WOM early on, preferably beginning in the prerelease period. After WOM is created, a high volume of WOM, which is beneficial to sales, tends to positively affect the volume of WOM in subsequent periods.

Summary, Managerial Implications, and General Discussion

This study contributes to the movie-marketing and WOM literature by examining movie WOM. Among other things, the mechanism through which WOM functions and the extent to which WOM helps explain box office sales are investigated. The data and analyses show that movie WOM is active during a movie's prerelease period and that WOM is more of a complement to other information sources than a substitute. Movie audiences tend to hold comparatively high expectation before a movie is released but become more critical in the opening week. The volume of WOM (but not the valence) offers significant explanatory power for box office revenue, both in the aggregate and for the early weeks, in support of a theory that WOM functions in the movie market primarily through an informative effect on awareness.

---

[13] Although the existence of movie stars and the measures of critical reviews do not significantly predict WOM, there are caveats regarding the roles of stars and movie critics. For example, some movie stars and critics may be on television or radio shows to talk about a to-be-released movie, which may lead to greater WOM. Furthermore, some stars may have popular Web sites, which also may influence WOM. An indicator variable of star power, albeit a typical approach in the literature, may not completely capture the effects of these scenarios (e.g., whether a particular movie has been promoted by its stars on popular television shows).

The finding that WOM is very active during the prerelease period implies that at least for entertainment goods, such as movies, WOM communication about a to-be-released new product may not depend on actual experience. Many potential users talk about the product on the basis of speculations. In this sense, managing buzz should be treated as a multidimensional task that extends beyond the new product itself. That is, management needs to be conscious about additional things that influence speculations, such as the firm's other products (both new and extant), the firm's reputation, various events, and possibly "rumors" that occur during the new product development process.

These findings highlight the need for studio and theater managers to observe and respond to WOM communications actively, especially during a movie's early weeks when most of the box office revenue is generated. However, what ultimately matters is the volume of WOM. It may not be critical to track information similar to the valence measures. Practically, volume can be measured either directly by a precise number, as this study does, or indirectly by proxies, such as the degree of public interest a movie receives, how widespread the communication occurs across demographic groups, and so on. These findings offer considerable support to the idea that firms should try to create active WOM communication among potential users (Godes and Mayzlin 2004a; Rosen 2000).

Another important managerial value of these findings is that WOM can be a useful measure to help forecast box office sales. Among other things, good forecasts help studios and theaters better plan screening capacity and potentially optimize exhibition contracts. However, it is not the purpose of this study to develop a full forecasting system. Rather, this study can be used to illustrate how the performance of a forecasting model can be improved by including actual WOM information. To do so, this study used the set of independent variables and adopted the standard calibrating-forecasting approach. Because the sample size makes it difficult to split the data into an estimation sample and a prediction sample, a rotating method was used to conduct the calibration and forecasting tasks. Specifically, each of the 40 movies is taken in turn as the movie to be predicted, while the other 39 movies are used in the calibration process to generate parameter estimates. The absolute percentage errors of the 40 predictions are then averaged as the indication of prediction effectiveness. The calibration prediction is conducted for both aggregate box office revenue and the opening-week box office revenue. The WOM measures are included in the proposed model but not in the benchmark model. Other independent variables are the same as those that appear in Tables 4 and 5.

This analysis shows that including WOM in the forecasting system significantly reduces forecasting errors by 31% for a movie's opening week (from a 55% absolute forecasting error with the benchmark model to 38% with the proposed model) and 23% for the aggregate box office revenue (from 61% to 47%). Different from most extant studies, these forecasts are made purely on the basis of the prerelease information of a new movie. This is probably the most difficult forecast task for movies (Sawhney and Eliashberg 1996). It is similar to Sawhney and Eliashberg's (1996) BOXMOD-I model when no revenue data for the new movie are available and a meta-analysis procedure is employed. Sawhney and Eliashberg report an average 71.1% prediction error for the aggregate box office revenue. Although the actual improvement through incorporating WOM depends on what forecasting system is eventually used, the added value of WOM appears to be significant.

As in previous studies (Basuroy, Chatterjee, and Ravid 2003; Eliashberg and Shugan 1997; Godes and Mayzlin 2004b), the current analysis focuses on statistical correlations between variables and uses these temporal relationships to infer how WOM may influence box office sales. Although these correlations offer interesting insights and useful managerial implications, further research might investigate the degree of causality, perhaps with more disaggregated data and different methodologies. In this regard, lab experiments could be a useful approach to examine the relationships such as those that Figure 2 illustrates.

Online user-generated information offers unique opportunities for research in marketing. Such data come in different formats, which render them suitable for different research purposes. Some data, such as the WOM messages used herein, are actual individual postings. Although their unstructured nature provides great flexibility for the analysis, coding them can be a difficult task. Other data, such as online user ratings, contain less flexibility but make it easier to conduct large-scale cross-sectional studies. Another issue involved with this growing literature is how similar the online information is to what occurs in the physical world. There is evidence that the degree of such similarity can be very high (Dellarocas, Awad, and Zhang 2004). Nevertheless, further research examining the difference between online user data and that in the physical context appears to have potential.

REFERENCES

Ackman, Dan (2001a), "Disaster of the Day," *Forbes*, (January 23), (accessed March 21, 2006), [available at http://www.forbes.com/2001/01/23/0123disaster.html].

——— (2001b), "Movie Theaters of the Absurd," *Forbes*, (March 2), (accessed March 21, 2006), [available at http://www forbes.com/2001/03/02/0302movies.html].

Ajzen, Icek and Martin Fishbein (1980), *Understanding Attitudes and Predicting Social Behavior.* Englewood Cliffs, NJ: Prentice Hall.

Anderson, Eugene (1973), "Consumer Dissatisfaction· The Effect of Disconfirmed Expectancy on Perceived Product Performance," *Journal of Marketing Research,* 10 (February), 38–44.

——— (1998), "Customer Satisfaction and Word of Mouth," *Journal of Service Research,* 1 (1), 5–17.

——— and Linda Salisbury (2003), "The Formation of Market-Level Expectations and Its Covariates," *Journal of Consumer Research,* 30 (June), 115–24.

Austin, Bruce (1989), *Immediate Seating: A Look at Movie Audiences*. Belmont, CA: Wadsworth.

Banerjee, Abhijit (1992), "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107 (3), 797–817.

Basuroy, Suman, Subimal Chatterjee, and S. Ravid (2003), "How Critical Are Critical Reviews? The Box Office Effects of Film Critics, Star Power, and Budgets," *Journal of Marketing*, 67 (October), 103–117.

Bayus, Berry (1985), "Word of Mouth: The Indirect Effect of Marketing Efforts," *Journal of Advertising Research*, 25 (3), 31–39.

Bowman, D. and D. Narayandas (2001), "Managing Customer-Initiated Contacts with Manufactures: The Impact on Share of Category Requirements and Word-of-Mouth Behavior," *Journal of Marketing Research*, 38 (August), 281–97.

Bristor, Julia (1990), "Enhanced Explanation of Word of Mouth Communications: The Power of Relationships," in *Research in Consumer Behavior*, E.C. Hirschman, ed. Greenwich, CT: JAI Press, 51–83.

Brown, Johnson and Peter Reingen (1987), "Social Ties and Word-of-Mouth Referral Behavior," *Journal of Consumer Research*, 14 (December), 350–62.

Burzynski, Michael and Dewey Bayer (1977), "The Effect of Positive and Negative Prior Information on Motion Picture Appreciation," *Journal of Social Psychology*, 101, 215–18.

Chaffee, Steven (1982), "Mass Media and Interpersonal Channels: Competitive, Convergent, or Complementary?" in *Inter/Media: Interpersonal Communication in a Media World*, Gary Gumpert and Robert Cathcart, eds. New York: Oxford University Press.

Dellarocas, Chrysanthos, Neveen Awad, and Michael Zhang (2004), "Exploring the Value of Online Reviews to Organizations: Implications for Revenue Forecasting and Planning," working paper, MIT Sloan School of Management

De Vany, Arthur and David Walls (1999), "Uncertainty in the Movie Industry: Does Star Power Reduce the Terror of the Box Office?" *Journal of Cultural Economics*, 23 (4), 285–318.

Duan, Wenjing, Bin Gu, and Andrew Whinston (2005), "Do Online Reviews Matter? An Empirical Investigation of Panel Data," working paper, Department of Management Science and Information Systems, University of Texas at Austin.

Eagly, Alice, Wendy Wood, and Shelly Chaiken (1978), "Causal Inferences about Communicators and Their Effects for Opinion Change," *Journal of Personality and Social Psychology*, 36 (April), 425–35.

Elberse, Anita and Jehoshua Eliashberg (2003), "Demand and Supply Dynamics for Sequentially Released Products in International Markets: The Case of Motion Pictures," *Marketing Science*, 22 (3), 329–54.

Eliashberg, Jehoshua, Jedid-Hah Jonker, Mohanbir Sawhney, and Berend Wierenga (2000), "MOVIEMOD: An Implementable Decision-Support System for Prerelease Market Evaluation of Motion Pictures," *Marketing Science*, 19 (3), 226–43.

——— and Steven Shugan (1997), "Film Critics: Influencers or Predictors?" *Journal of Marketing*, 61 (April), 68–78.

Faber, Ronald and Thomas O'Guinn (1984), "Effect of Media Advertising and Other Sources on Movie Selection," *Journalism Quarterly*, 61, 371–77.

Godes, David and Dina Mayzlin (2004a), "Firm-Created Word-of-Mouth Communication: A Field-Based Quasi-Experiment," Harvard Business School Working Paper No 05-023.

——— and ——— (2004b), "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science*, 23 (4), 545–60.

Graser, Marc (2002), "Semel Gives Yahoo! the Business: Internet Portal Ratchets Up H'wood Links," *Variety*, (August 19–25), 11.

Harrison-Walker, L. Jean (2001), "The Measurement of Word-of-Mouth Communication and an Investigation of Service Quality and Customer Commitment as Potential Antecedents," *Journal of Services Marketing*, 4 (1), 60–75.

Holbrook, Morris (1999), "Popular Appeal Versus Expert Judgments of Motion Pictures," *Journal of Consumer Research*, 26 (September), 144–55.

King, Tom (2001), "Will 'Harry' Have Legs? After Record Opening, Wizard Falls Off at the Box Office; Trying to Catch R2-D2," *The Wall Street Journal*, (November 30), W4.

——— (2003), "Now Playing Somewhere Else: Studios Try to Build Buzz by Holding Films Back; in Muncie? Read a Book," *The Wall Street Journal*, (January 10), W10.

Krider, Robert and Charles Weinberg (1998), "Competitive Dynamics and the Introduction of New Products: The Motion Picture Timing Game," *Journal of Marketing Research*, 35 (February), 1–15.

Lehmann, Donald and Charles Weinberg (2000), "Sales Through Sequential Distribution Channels: An Application to Movies and Videos," *Journal of Marketing*, 64 (July), 18–33.

Lieberman, Al and Patricia Esgate (2002), *The Entertainment Marketing Revolution: Bring the Mogul, the Media, and the Magic to the World*. Upper Saddle River, NJ: Financial Times/Prentice Hall

Litman, Barry (1983), "Predicting Success of Theatrical Movies: An Empirical Study," *Journal of Popular Culture*, 16 (Spring), 159–75.

Litwak, Mark (1986), *Reel Power: The Struggle for Influence and Success in the New Hollywood* New York: William Morrow.

Mahajan, Vijay, Eitan Muller, and Roger Kerin (1984), "Introduction Strategy for New Products with Positive and Negative Word-of-Mouth," *Management Science*, 30 (12), 1389–1404.

Mizerski, Richard (1982), "An Attribution Explanation of the Disproportionate Influence of Unfavorable Information," *Journal of Consumer Research*, 9 (3), 301–310.

Motion Picture Association (2003), *U.S. Entertainment Industry. 2003 MPA Market Statistics*. Encino, CA: Motion Picture Association.

Murray, Keith (1991), "A Test of Services Marketing Theory: Consumer Information Acquisition Activities," *Journal of Marketing*, 55 (January), 10–25.

Neelamegham, Ramya and Pradeep Chintagunta (1999), "A Bayesian Model to Forecast New Product Performance in Domestic and International Markets," *Marketing Science*, 18 (2), 115–36.

Oliver, Richard (1980), "Conceptualization and Measurement of Disconfirmation Perceptions in the Prediction of Consumer Satisfaction," in *Proceedings of Fourth Annual Conference on Consumer Satisfaction, Dissatisfaction, and Complaining Behavior*, H. Hunt and R. Day, eds. Bloomington: Indiana University School of Business.

Olson, J and P. Dover (1979), "Disconfirmation of Consumer Expectation Through Product Trial," *Journal of Applied Psychology*, 64 (April), 179–89.

Petty, Richard and John Cacioppo (1984), "The Effects of Involvement on Responses to Argument Quantity and Quality: Central and Peripheral Routes to Persuasion," *Journal of Personality and Social Psychology*, 46 (January), 69–81.

Reddy, Srinivas, Vanitha Swaminathan, and Carol Motley (1998), "Examining the Determinants of Broadway Show Success," *Journal of Marketing Research*, 35 (August), 370–83.

Reinstein, David and Christopher Snyder (2005), "The Influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics," *Journal of Industrial Economics*, 53 (1), 27–51.

Rogers, Everett (1983), *Diffusion of Innovations*. New York: The Free Press.

Rosen, Emanuel (2000), *The Anatomy of Buzz: How to Create Word-of-Mouth Marketing*. New York: Doubleday.

Russell, Laczniak, Thomas DeCarlo, and Sridhar Ramaswami (2001), "Consumers' Responses to Negative Word-of-Mouth Communication: An Attribution Theory Perspective," *Journal of Consumer Psychology*, 11 (1), 57–73.

Sawhney, Mohanbir S. and Jehoshua Eliashberg (1996), "A Parsimonious Model for Forecasting Gross Box-Office Revenues of Motion Pictures," *Marketing Science*, 15 (2), 113–31.

Schiffman, Betsy (2001), "Movie Theaters Stuck in 'Stone Age," *Forbes*, (September 22), (accessed March 21, 2006), [available at http://www.forbes.com/2000/09/23/0923newmedia.html].

Shugan, Steven (1995), *Service Strategy*. Homewood, IL: Richard D. Irwin.

Swami, Sanjeev, Jehoshua Eliashberg, and Charles Weinberg (1999), "SilverScreener: A Modeling Approach to Movie Screens Management," *Marketing Science*, 18 (3), 352–72.

Taylor, John (2003), "Word of Mouth Is Where It's At," *Brandweek*, (June 2), 26.

Van den Bulte, Christophe and Gary Lilien (2001), "Two-Stage Partial Observability Models for Innovation Adoption," working paper, Wharton School, University of Pennsylvania.

Vogel, Harold (2001), *Entertainment Industry Economics: A Guide for Financial Analysis*, 5th ed. Cambridge, UK: Cambridge University Press.

Zufryden, Fred (1996), "Linking Advertising to Box Office Performance of New Film Releases: A Marketing Planning Approach," *Journal of Advertising Research*, 36 (July–August), 29–41.

Jan B. Heide & Kenneth H. Wathne

Friends, Businesspeople, and Relationship Roles: A Conceptual Framework and a Research Agenda

This article develops a conceptual framework of roles in marketing relationships. Drawing on emerging theory from economic sociology and March's (1994) notion of decision "logics," the authors discuss two prototypical relationship roles, namely (1) a "friend," who uses a "logic of appropriateness" and follows established rules, and (2) a "businessperson," whose decisions are guided by utility-maximizing considerations under a "logic of consequences." Next, they use extant theories of interfirm governance to suggest that firms' relationship strategies can be used both to create different relationship roles in the first place and to activate them over time. The authors posit that activation can have several different outcomes, including reinforcement of an existing dominant role or actual switching to a new one. Theoretically, the conceptual framework allows for integration of different perspectives on interfirm relationships, some of which have provided seemingly inconsistent accounts of firm behavior. From a managerial perspective, the framework identifies specific matches between particular relationship roles on the one hand and firms' governance strategies on the other hand. In general, the framework suggests that an in-depth understanding of roles is a prerequisite for the deployment of relationship management initiatives toward resellers, customers, and suppliers.

A considerable body of literature exists on the general orientations that exchange partners bring to bear on their relationships. In branches of the new institutional economics literature, decision makers' assumed principle of action is utility maximization, sometimes to the point of "self-interest seeking with guile" (e.g., Williamson 1985, p. 47). This assumption has produced an extensive body of research on so-called governance mechanisms (e.g., monitoring and incentive devices), which have the presumed ability to manage parties' inherent "calculative" orientations (see, e.g., Rindfleisch and Heide 1997). In contrast, other streams of literature, particularly in sociology, have been overtly critical of such an "undersocialized" view of exchange, arguing that relationship behavior follows from rules or "heuristics," which are applied regardless of their economic consequences (e.g., Granovetter 1985; Uzzi 1997).

Notably, although these economic and sociological perspectives provide different accounts of relationship behavior, they actually *share* a fundamental underlying assumption, namely, that exchange partners are "unitary actors" who play a single role over time and consistently make decisions according to a fixed pattern (i.e., in a calculative or heuristic manner). However, this assumption has been challenged. For example, Montgomery (1998) argues that attributing global characteristics to exchange partners, either calculative or heuristic, is potentially misleading. Using a "prisoner's dilemma" game as an illustration, he suggests that the "players" are not individuals with fixed traits or predispositions, as is often assumed, but rather collections of roles. In his analysis, Montgomery distinguishes between two role archetypes: (1) a "businessperson," whose decisions are guided by utility maximization considerations and who will "defect" if such a move is individually profitable, and (2) a "friend," who will cooperate as a matter of principle.[1]

In this article, we build on this general view of relationship parties and interactions to develop a conceptual framework of roles in marketing relationships. Our framework explicitly accounts for the possibility that different orientations (calculative and heuristic) and the associated roles (businessperson and friend) can coexist and that switching among them is both possible and likely. Moreover, we argue

Jan B. Heide is Irwin Maier Chair in Marketing, School of Business, University of Wisconsin, Madison, and Professorial Fellow, Department of Management, University of Melbourne (e-mail: jheide@bus.wisc.edu). Kenneth H. Wathne is Assistant Professor of Marketing, School of Business, University of Wisconsin, Madison (e-mail: kwathne@bus.wisc.edu). Part of this article was completed while Heide was Montezemolo Visiting Professor in the Judge Business School, University of Cambridge. The authors thank Aric Rindfleisch and Maria Papas Heide for their helpful comments They also thank the anonymous *JM* reviewers for their constructive suggestions. The authors are listed in alphabetical order. Each author contributed equally to the article.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

---

[1] We acknowledge that these particular labels carry surplus meaning and can be misleading. Any given friend is unlikely to be devoid of task orientation, and real-life businesspeople frequently show exemplary service. Nevertheless, we rely on these particular terms for two reasons: First, they have been used in the literature (e.g., Montgomery 1998) as "ideal type" (in a Weberian sense) theoretical constructions. Second, field studies have shown that these specific labels are used in actual industry settings.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 70 (July 2006), 90–103

that firms' governance strategies (e.g., socialization and monitoring) are capable of (1) creating relationship roles, (2) reinforcing existing roles over time, and (3) producing role switching (e.g., from friend to businessperson).

From a theoretical standpoint, our framework provides a novel perspective on relationship management. As we discuss subsequently, our perspective can explain empirical results that run counter to established governance theory; for example, certain governance mechanisms actually promote rather than suppress opportunistic behavior (e.g., John 1984).

The role concept also has important practical implications. Ridgway (1957, p. 465) noted this early on, claiming that "the establishment and maintenance of the manufacturer-dealer system itself—that is, the provision of mutually satisfying roles in the relationship—is a major administrative responsibility." Importantly, although industry observers have since made a case for the importance of relationship roles (e.g., Normann and Ramirez 1993), specific guidelines for decision making are almost absent from the literature. Frazier (1999, p. 235) has argued that (1) "little is known about the specification of channel roles" and (2) "many manufacturers make role decisions very poorly."

We organize the article as follows: We provide a brief synthesis of existing theoretical perspectives and propose a new decision-theoretic conceptualization of relationship roles. Then, we present our conceptual framework, discuss the implications, and suggest areas for further research.

Role Theory: Origins and Emergent Views

According to Biddle (1986, p. 68), role theory began as a "theatrical metaphor." Over time, the general domain of role theory has become the "study of behaviors that are characteristic of actors within contexts" (Biddle 1979, p. 56). Although several different approaches to role theory exist (e.g., Kahn et al. 1964; Mead 1932; Sherif 1936), the two branches, known as functional (e.g., Bates and Harvey 1975; Parsons 1937) and structural (e.g., Burt 1982; Winship and Mandel 1983) role theory, have historically enjoyed some prominence and have also received attention in the marketing literature. For example, applications of functional role theory in marketing are evident in the use of terms such as "wholesaler" and "retailer" to describe particular channel positions with associated functions and activities (e.g., El-Ansary and Stern 1972). Structural role theory has been used in marketing to describe the properties of distribution networks and their interaction patterns (e.g., Anderson, Håkansson, and Johanson 1994; Iacobucci 1996). In general, however, despite Stern's (1969) early call, role theory has not been extensively used in research on interfirm phenomena.[2]

Although conventional role theory has generated important insights, it has also been criticized (e.g., Powell and Smith-Doerr 1994; Salancik 1995) for, among other reasons, not going far enough in terms of articulating the properties of the relevant parties and their decision-making patterns. For example, Blau (1993) notes that a party's actual behavior in a relationship is not necessarily isomorphic with its position. This is consistent with Larson's (1992) observation that standard role designations, such as "principal" and "agent," often provide little information about the actual properties of a set of parties and their relationship.

Emerging role theory has attempted to address these limitations. In this article, we draw on recent role research in economic sociology (e.g., Montgomery 1998) and define an organizational role as an organizational "identity" or a "collective mind" (Kohli and Jaworski 1990), which provides the foundation for shared perceptions and coordinated decision making (Albert and Whetten 1985; Messick 1999; Tajfel and Turner 1986). Our focus is on two particular archetypes of roles or identities—friend and businessperson—which differ systematically in their decision-making patterns.

Role of a Friend: The Logic of Appropriateness

Most versions of role theory presume that expectations, broadly defined, are major determinants of role behavior (Biddle 1986). The term "expectation" is commonly used in a prescriptive sense, referring to behaviors that "ought to" be performed (Turner 1974). In certain branches of role theory (e.g., Montgomery 1998), the term "rule" is used to describe how expectations guide organizational decision making.

Montgomery's (1998) role-theoretic arguments are derived from the work of March (1994), who argues that decision makers explicitly or implicitly ask three key basic questions. These questions pertain to (1) recognition ("What kind of situation is this?"), (2) identity ("What kind of organization is this?"), and (3) rules ("What does an organization such as this do in a situation such as this?"). Conceptually, the first two questions involve categorization tasks (Turner et al. 1987)—for example, the categorization of a firm's exchange partner and the firm itself—and the third question involves the matching of behaviors to categories (Forgas 1982). For example, a distributor that agrees to carry additional inventory for a manufacturer may be categorized as "cooperative" and, subsequently, may be granted "gold partner" status.

Research in contract law and marketing (e.g., Dwyer, Schurr, and Oh 1987; Heide and John 1992; Macaulay 1963) describes how rules of behavior in interfirm relationships may manifest themselves as norms. For example, a supplier and a buyer may develop a norm of flexibility, which specifies how certain aspects of the relationship will be modified under changing circumstances. March (1994) uses the term "logic of appropriateness" to describe such a rule-driven decision-making mode.

Montgomery (1998) uses March's (1994) logic of appropriateness to coin a prototypical relationship role, namely, that of a friend. Using a game-theoretic analogy, Montgomery defines a friend role in terms of a tendency to play a game based on fixed rules that prescribe cooperative moves, even when such moves undermine a party's individual payoff. Uzzi's (1997) study provides several "thick descriptions" of such a role. For example, a production manager in an apparel company described how her firm would forgo self-gain by contracting early for services to help a friend suffering from slow business. In the process, the company purposely sacrificed potential profits from alternative contractor relationships.

---

[2] Role theory has been used more extensively in research on intraorganizational phenomena (e.g., Behrman and Perreault 1984; Churchill et al. 1985; Singh 2000; Webster and Wind 1972).

Role of a Businessperson: The Logic of Consequences

Although rules may serve as explanations for behavior, and roles themselves (e.g., a friend) can be conceptualized on the basis of rule-following behavior, some researchers suggest that roles are induced through the consequences of behavior (Biddle 1979). This principle is the cornerstone of rational choice theory (e.g., Becker 1976; Hogarth and Reder 1987), which tends to describe decision-making processes in utility maximization terms. March (1994, 1999) refers to such processes as involving a "logic of consequences."

Many of the extant theories of interfirm relationships, including transaction cost, agency, and game theory, share the idea of anticipatory, consequential choice. For example, a common explanation for sustained cooperation between firms is "calculative trust" (Williamson 1993).[3] For example, a manufacturer may trust a distributor to fulfill its role because the known short-term benefits to the distributor from opportunism are outweighed by the long-term benefits from cooperation.

The prototypical relationship role that corresponds to a logic of consequences is that of a businessperson, who is not bound by rules per se but whose primary motivation is utility maximization. In the terminology of game theory (Axelrod 1984), a businessperson plays the game by choosing the strategy (cooperation or defection) that maximizes individual payoffs. A businessperson may even pursue opportunistic behaviors (e.g., quality shirking) at the expense of a partner, if the gains from doing so are sufficiently large (Klein 1996).

Distinctions Between the Two Logics

Importantly, although rational choice theory does not deny the importance of rules, it tends to view them as being derived from, or somehow supported by, underlying utility maximization considerations (Scott 2000). For example, although Gibbons (1999) recognizes the importance of informal rules within relational contracts, he also notes how such rules must be explicitly enforced in a relationship (e.g., through a pattern of repeated interaction). As such, rule-following behavior (and, indeed, the very notion of relational contracting) is inherently tied to (economic) consequences.

However, other scholars argue that the standard rational choice characterization of utility maximization is incomplete. For example, Montgomery (1998, p. 100) argues,

> Rational choice theorists have trouble explaining adherence to "social norms" not because these rules would be difficult to derive from utility maximization, but because the associated maximization problems (entailing merely "psychic" costs and benefits) often seem vacuous. The need to construct nontrivial utility-maximization problems have led rational choice theorists to emphasize external sanctions (I follow a norm because others are watching me) over internalization (I follow a norm because of who I am). To the extent that a role has been internalized (i.e., the individual's self-assessed "degree of membership" is positive), the cost of disobeying the norm is found not merely in external sanctions (which are completely irrelevant if the individual knows that she is not being monitored), but in the violation of self-consistency—individuals would no longer recognize themselves.

Montgomery's account of an "appropriateness" mode of decision making is consistent with the "moral imperative" view that Durkheim (1915), Etzioni (1975), and Elster (1989) describe. For example, Etzioni argues (p. 24) that "utilitarian satisfaction should not be confused with the sense of affirmation that accompanies discharging one's moral commitments, commitments that are often in themselves taxing rather than pleasurable."[4] Moreover, Elster writes (p. 104), "When norms are internalized, they are followed even when violation would be unobserved and not exposed to sanctions. People have an internal gyroscope that keeps them adhering steadily to norms, independently of current reactions of others." On this note, and in contrast with Gibbons's (1999) view, Hackett (1994) provides evidence of "pure" relational contracts, in which the rules themselves are capable of guiding decision making, even in the absence of explicit enforcement devices.

On the basis of the previous discussion, we suggest that two distinct decision logics exist and that they can be used to delineate two prototypical relationship roles, namely, friend and businessperson. We summarize the preceding discussion in Table 1.

---

[3] Barney and Hansen (1994) make a distinction between "semi-strong" and "strong" or "hardcore" forms of trust. Semistrong trust corresponds to Williamson's (1993) and Uzzi's (1997) notion of calculative trust and reflects a calculative decision logic in line with the work of March (1994). Conversely, strong trust corresponds to Uzzi's (1997) notion of heuristic trust and reflects a rule-based decision logic in line with the work of March (1994). As Barney and Hansen note (p 179), hardcore trustworthy exchange partners are trustworthy because "that is who they are."

[4] Work within the rational choice paradigm has tried to reconcile these perspectives by suggesting that decision makers possess multiple utility functions for individual and social welfare, respectively (e.g., Harsanyi 1955; Margolis 1982). As such, utility may exist at different *levels*. In the marketing literature, Rokkan, Heide, and Wathne (2003) describe how solidarity norms prescribe behaviors that maximize relationship utility. In addition, it has been argued that utility exists in different *forms*, including that of psychological "warm glow" (Andreoni 1989, p. 1448). Still, the assumption that decision makers engage in calculations to maximize utility (conceptually, a logic of consequences) is consistent across all these models.

TABLE 1
Prototypical Relationship Roles

	Role	
Key Characteristics	**Friend**	**Businessperson**
General theoretical origin of role	Sociology	Economics/rational choice theory
Subfields	Economic sociology (embeddedness perspective) Pure (self-sustaining) relational contracts	Transaction cost theory Agency theory Game theory Externally enforced relational contracts
Representative literature	Granovetter 1985; Etzioni 1975; Elster 1989; Macaulay 1969	Williamson 1985, 1996; Jensen and Meckling 1976; Axelrod 1984; Gibbons 1999
Role's underlying decision logic[a]	Appropriateness (heuristic)	Consequences (calculative)
Role's guide to decision making[b]	Rules	Incentives
Actual role behavior	Cooperation in principle	Defection or cooperation = f (incentive structure)

[a] See March (1994).
[b] See Montgomery (1998).

The Coexistence of Logics: A Decision-Theoretic Perspective on Relationship Roles

Notably, although the bodies of theory that provide the basis for the two prototypical relationship roles are quite different, as Table 1 shows, they tend to converge on the same general assumption, namely, that decision makers either follow rules or maximize utility or, more generally, that they play a single role (friend *or* businessperson) over time. In Gambetta's (1996) terminology, a decision maker either (1) "is pushed from behind" (by preexisting rules) or (2) "jumps" (toward the alternative with the highest payoff).

We wish to challenge this assumption. Although we view the two roles as distinct, in the sense that one or the other may be applied to an individual decision, we believe that the two can also coexist within a given party in the form of latent predispositions (Friedberg 2000; Greenwald and Pratkanis 1984). This is consistent with Powell and Smith-Doerr's (1994, p. 371) argument that "identities are forged out of porous, multi-stranded relations in which business, reputation, and friendship are entangled."

A largely unaddressed issue, however, pertains to the specific manner in which roles coexist. Consider again the argument that an organizational role represents a form of identity or mental model of a relationship (Albert and Whetten 1985; Messick 1999). At a given time, based in part on a relationship's unique interaction history, a particular identity will dominate a party's approach to that relationship (Oakes 1987). For example, a manufacturer may view a certain supplier as a friend, and the manufacturer's general approach to the relationship may be reflected in friend types of actions, as in Uzzi's (1997) example, which we described previously.

However, although the supplier in question may have friend status, the manufacturer can simultaneously possess a mental model of businessperson suppliers based on actual experiences in other relationships or on general industry observation. Over time, depending on the focal supplier's actions, the established relationship roles may be subject to reassessment.

Research in organization theory (e.g., Albert and Whetten 1985) describes how identities are formed by interorganizational comparison processes that directly affect organizational action. If a given comparison process produces a sufficiently large "identity gap," certain congruence-enhancing responses are triggered, such as a reassessment of a party's own identity and/or the partner's. For example, to the extent that a supplier pursues actions normally associated with a businessperson role, a complex process of role switching may be initiated. For example, the former supplier friend may be assigned to a new category (a businessperson), and the manufacturer's own role may receive a similar designation.

Larson (1992, pp. 86–100) provides an example of how a role shift can also be triggered by external factors:

> [T]he role shift occurred when "Support Products" (SP) was acquired by a much larger corporation. Before it was acquired, the company had a reputation for developing close supplier relationships, characterized by an ethos of friendship and mutual assistance. SP's procurement manager described the philosophy behind the firm's key supplier partnerships: "These people are extensions of your own organization—they are part of your family." When SP was acquired, the new parent imposed internal sourcing requirements on its newly acquired division that dramatically altered the terms and character of SP's supplier alliances. The partnership philosophy was replaced by a short-term, cost-only focus. SP's purchasing manager explained that as a consequence, the relationships shifted,

and collaboration and special treatment became increasingly unlikely.

At a general level, role-switching processes have been described as transitions between "primary and secondary frameworks" (Goffman 1969) or between "quasi-stationary equilibria" (Ashforth 2001). The actual transitions have sometimes been described in terms of "unfreezing" and "refreezing," suggesting that role switching is subject to inertia pertaining to both (1) a partner's new designation and (2) a party's own role and associated behaviors. Regarding the former, a given event or situational cue (e.g., a supplier that used to be viewed as a friend refuses to adapt to changing circumstances) may trigger doubts about the viability of the existing role. However, an actual recategorization of the supplier need not take place, at least not without additional confirmatory evidence. A chief executive officer who Uzzi (1997, pp. 44–45) interviewed provided the following description: "Sometimes they ask for a favor (e.g., a low price) and we'll do it. But if they always do that, they're ripping me off."

Similarly, the focal party (i.e., the manufacturer) may continue to play an existing role, even in the face of evidence about a partner's new role, to the extent that the focal actions reaffirm the firm's own identity and, in general, reduce uncertainty (Turner et al. 1987). In addition, an existing role may be "sticky" (Bolton and Reed 2004) to the extent that it is shared by other parties or referents (Ashforth 2001). For example, a manufacturer that has a global reputation as a friend in a larger industrial network may continue to enact a friend role, even toward suppliers whose actions belong to a standard businessperson repertoire. Decision makers may have a baseline tendency to pursue behaviors that reaffirm existing roles.

Numerous accounts of firms' struggles to change their relationship orientations provide evidence of role stickiness. For example, Lyons, Krachenberg, and Henke (1990, p. 33) report the following statement that a senior executive made about his firm's attempted change in supplier strategy: "We've had to outplace or retire some of our most experienced, veteran buyers. It was just too much to expect them to change from playing poker with suppliers to cooperating with them. The old ways and the new game just didn't match."

Despite the possibility of role stickiness, the general conclusion from the previous discussion is that parties may possess multiple identities over time, even within the context of an individual relationship. As such, rather than applying fixed characterizations, such as calculative or heuristic, we view relationship parties as collections of roles. Although the dual role conceptualization of friend and businessperson is admittedly somewhat simplistic, it nonetheless provides a useful starting point for building a theory of relationship roles.

Roles and Governance Decisions: A Conceptual Framework

In this section, we outline a framework that specifies how roles are (1) initially created in relationships and (2) subsequently activated. We also sketch out possible linkages between roles and particular relationship outcomes. As we discuss throughout the section, these outcomes pertain both to the roles themselves (e.g., reinforcement of current roles or switching) and to various economic dimensions. For example, the consequence of a firm's decision to begin monitoring a supplier "friend," which, by definition, would not require external surveillance, may be an economic opportunity cost.

In developing the framework, we integrate the notion of a role with extant frameworks of interfirm governance (e.g., Cannon and Perreault 1999; Heide 1994; Ouchi 1979; Wathne and Heide 2000). The governance literature provides a useful building block because it identifies particular relationship management mechanisms with different time dimensions, which can be expected to serve particular role creation and activation purposes. We first review a core set of governance mechanisms and their general properties. Then, we discuss their particular purposes in relation to relationship roles.

General Governance Mechanisms

Ouchi (1979) discusses two broad relationship management strategies available to firms: A firm can (1) identify parties that "fit its needs exactly" (p. 840) or (2) design a managerial system that specifically promotes desired behaviors. In practice, as suggested by extant governance typologies (e.g., Heide 1994; Wathne and Heide 2000), the former approach can be implemented through selection and socialization efforts, whereas the latter tends to be based on incentives and monitoring. Although our primary focus is not on these mechanisms per se, a brief review of each, including examples of actual industry practice, provides a useful foundation for the presentation of our role framework.

Selection. Relationship governance through selection involves identifying exchange partners a priori that exhibit a good fit on particular criteria. The selection criteria may involve both skills (e.g., product quality and human resources) and values (e.g., cultural compatibility). Recent industry accounts (Killian 2003) show how companies such as Culver's and Home Depot rely on both types of criteria to select their franchisees and suppliers, respectively. Ford Motor Company has formalized its selection efforts through specialized supplier qualification programs, which serve as a trial period for prospective suppliers (*Purchasing* 1995). For some companies, such as the food retailers Whole Foods Market and Pret A Manger, selection is such an integral part of the companies' overall strategies that their elaborate practices are explicitly communicated to their stakeholders. For example, Pret A Manger's promotional materials describe its selection practices as "treasure hunting."

Socialization. Conceptually, socialization is the process by which new parties learn skills and internalize another party's values, goals, and rules (Chatman 1991). Socialization can happen informally over the normal course of a relationship, or processes can be explicitly designed to promote learning and internalization. As an example of the latter, many corporations have established their own universities as a way to both develop particular skills and promote particular values. These efforts range from McDonald's attempt to get franchisees to "March to a single McDrum" (*The Economist* 1999), to Saturn's investments in "Saturnizing" dealership employees (Meister 1998), to Motorola's training of supplier and distributor "Black Belts" (*Corporate University Review* 2000). Supplier qualification programs, as we discussed previously, may also serve distinct socialization purposes, in that a party is asked to take on a future role in an "anticipatory" way (Biddle 1979; Dubinsky et al. 1986).

Incentives. A cornerstone of agency theory (e.g., Jensen and Meckling 1976) is relationship management through explicit incentives. For example, manufacturers can manage their reseller relationships through commissions (John and Weitz 1989), margins (Dutta, Bergen, and John 1994), bonuses (Gilliland 2003), or some combination thereof. Cisco recently launched a reseller program called "Take Stock in Cisco," which provides "share points" in return for selling particular Cisco products (Schick 2001). In general, the rationale for the use of incentives is to structure a relationship in such a way that particular actions are explicitly rewarded or punished (Telser 1980).

Monitoring. The final governance strategy involves explicit measurement of partner performance relative to some preexisting agreement or standard (Anderson and Oliver 1987). In principle, a firm's monitoring may include a partner's output or behavior, and it can be implemented in various ways. For example, industrial buyers may monitor their component suppliers using incoming inspection, evaluation of statistical process control charts, or actual audits of a supplier's plant. Firms such as Sears have formalized their monitoring process by developing "vendor report cards" (Aron 1998). Franchisors such as McDonald's rely on "undercover" agents who evaluate franchisees on attributes such as food quality, speed, and store cleanliness (Tatge 2001). Recently, Arby's and 7-Eleven have begun to experiment with electronic surveillance systems (Shirouzu and Bigness 1997).

Our conceptual framework (see Figure 1) is based on the premise of inherent matches between particular governance mechanisms on the one hand and different relationship roles on the other hand. Specifically, we propose that relationship strategies, which emphasize "clanlike" selection and socialization, are inherently linked with the role of a friend and rule-based decision making under a logic of appropriateness. In contrast, the "harder" governance mechanisms of incentives and monitoring are linked to a businessperson role and its logic of consequences. Our framework suggests that a firm's governance decisions can affect both (1) the creation of relationship roles *ex ante* and (2) the activation of roles over time.

Ex Ante *Role Creation*

In principle, all the governance mechanisms we discussed previously are capable of creating relationship roles. However, the specific manner in which the relevant roles are created may differ dramatically.

Consider selection and socialization as a starting point. Such efforts are inherently associated with a logic of appropriateness, in the sense that a firm attempts either (1) to identify a friend that already possesses the relevant rules or (2) to develop one through social processes. Supplier qualification programs, which we discussed previously, may create friend roles in both of these manners. Note that the role creation process under a qualification program is direct, in that one firm (e.g., a buyer) targets its governance efforts toward a particular party (e.g., an individual supplier). Theoretically, role creation in this manner involves creating an "embedded" market through social processes (Granovetter 1985). We illustrate the relevant processes in Figure 1 through the top-left vertical arrow.

Governance mechanisms may also establish roles in an indirect way by allowing parties to signal their predispositions and self-select into the relationship. Consider the previously discussed businessperson role and governance efforts in the form of incentives and monitoring. As we discuss in further detail subsequently, such strategies may be inherently incompatible with a friend role. However, to the extent that a given firm has a known reputation *ex ante* for relying on monitoring and incentives, it allows other parties, which already possess dominant businessperson roles and for which such mechanisms are consistent with their decision-making mode (i.e., a logic of consequences), to self-select into the relationship.

As Stryker (1968) describes, people have a tendency to seek out opportunities to enact a "salient identity." Gambetta (1996) provides an interesting account of how publicly known incentive systems serve as distinct self-selection devices into academic institutions. Similarly, franchisors attract franchisees to run their independent stores whose mind-sets are consistent with incentive systems based on "residual claims" (Brickley and Dark 1987). The Swedish furniture company IKEA has a well-known system of supplier competition, in which both incumbent and integrated suppliers must compete like independent contractors with new suppliers on an ongoing basis (Margonelli 2002). This system creates a particular form of incentive structure, which allows "appropriate" suppliers (i.e., those that possess compatible orientations) to self-select into the manufacturer relationship. The (indirect) creation of a businessperson role in this manner involves the establishment of a conventional "market" rather than an "embedded" relationship (Baker 1990; Williamson 1975). The top-right vertical arrow in Figure 1 depicts these processes.[5]

---

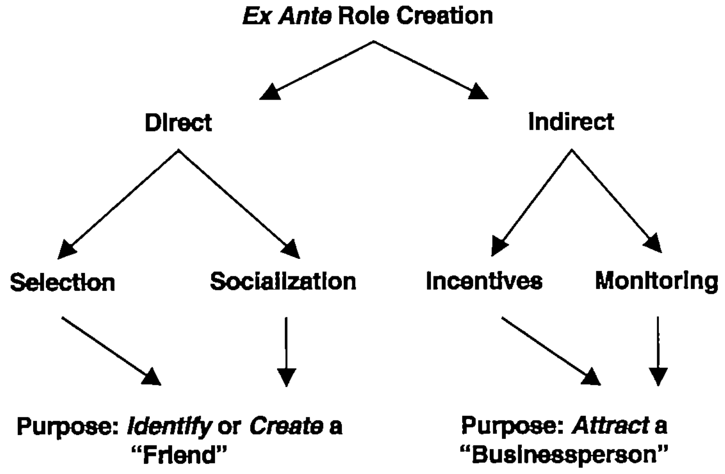
[5] For simplicity, in the previous discussion (and in Figure 1), we associated direct governance efforts with selection and socialization and a friend role, but firms may also rely on such efforts to create a businessperson. For example, Ryan and Deci (2000) describe how one party may internalize another's economic goals. Conceptually, however, the ongoing behavior that follows from this form of internalization is guided by utility-maximizing considerations under a logic of consequences. Moreover, although we discussed the possibility of (indirect) self-selection in the context of monitoring and incentives and a businessperson role, self-selection can also be associated with other governance mecha-

Ex Post *Role Activation*

In addition to creating roles *ex ante*, as we discussed previously, governance mechanisms also play important *ex post* functions in relation to relationship roles. As was the case previously, our framework is based on the assumption that certain linkages exist between (1) the decision "logics" that underlie particular roles and (2) the inherent nature of particular governance mechanisms.

These linkages may manifest themselves in two different ways: First, the role that currently dominates a party's approach to a certain relationship may be reinforced by the other party's governance choices. The reinforcement scenarios, which involve "matches" between roles and governance mechanisms, appear in Cells 1, 5, and 6 in Figure 1. Alternatively, to the extent that a particular governance

FIGURE 1
Conceptual Framework


Ex Ante Role Creation
Direct
Indirect
Selection
Socialization
Incentives
Monitoring
Purpose: *Identify* or *Create* a "Friend"
Purpose: *Attract* a "Businessperson"

Ex Post Role Activation[a]		Purpose: *Identify* or *Create* a "Friend"	Purpose: *Attract* a "Businessperson"
	Socialization	**1. Role Reinforcement** → Increased cooperation	**4. Role Maintenance** → Resource waste
	Incentives	**2. Role Dilution** ↓ •Perfunctory compliance •Opportunity costs → **Role Switching** ↓ •Retaliatory action •Direct transaction costs	**5. Role Reinforcement** → Increased cooperation
	Monitoring	**3. Role Dilution** ↓ •Perfunctory compliance •Opportunity costs → **Role Switching** ↓ •Retaliatory action •Direct transaction costs	**6. Role Reinforcement** → Increased cooperation

☐ Match scenario ▨ Mismatch scenario

[a]Unlike the other three governance mechanisms, by definition, selection is used only *ex ante*.

---

nisms (e.g., socialization). For example, Camerer and Vepsalainen (1988) discuss how companies that make their cultures visible create opportunities for self-selection.

mechanism is incompatible with the current role's decision logic, a mismatch scenario is created, as we show in Cells 2, 3 and 4. If the current role is sticky, as we previously discussed, the immediate outcome of a mismatch scenario may be maintenance or perhaps dilution of the current role. However, there is also the possibility that a mismatch will produce actual role switching.

As we discuss subsequently, the upshot of our framework is that the different *ex post* scenarios produce immediate outcomes with respect to the roles themselves (e.g., reinforcement or switching). In addition, the different scenarios have long-term consequences of both a behavioral and an economic nature. We first discuss the possible *ex post* scenarios for an existing friend role, and then we sketch out the parallel scenarios for a businessperson role.

Existing friend role. Assume that a firm's initial governance efforts toward a particular exchange partner involved selection or socialization and that these efforts succeeded in establishing a friend role for the partner. Given these initial efforts, the relationship will end up in Cell 1, Cell 2, or Cell 3, depending on the subsequent governance choices. If the firm continues to rely on socialization (Cell 1), the partner's current role will be reinforced. For example, a supplier that has achieved friend status may sustain and possibly increase the original level of rule-based cooperation. Given the evidence that initial socialization effects tend to decay over time (Harrison and Caroll 1991), this form of role reinforcement may be an important relationship initiative. As a specific example, Tupperware organizes regular "homecoming jubilees" for its top sellers, which include the singing of an anthem called "I've got that Tupper Feeling" (Brooks 2004).

However, if the firm in question deploys explicit incentives (Cell 2) or monitoring (Cell 3), there is a distinct possibility of a mismatch between (1) the decision logic inherent in the established role (i.e., a friend's logic of appropriateness) and (2) the use of governance mechanisms that are inherently tied to a logic of consequences. From a purely descriptive perspective, the available evidence suggests that such mismatches are not unlikely. For example, Gardiner (2003, p. 18) argues that firms "are not good at judging motivations" and, consequently, are prone to "heavy and misplaced" reliance on incentive-based strategies.

What are the potential outcomes of a mismatch? It is possible that the existing role will continue to be enacted by the partner, though in a weakened or diluted state. Specifically, a new governance regime may create an "identity gap," as we previously discussed, and trigger doubts about the viability of the existing relationship roles. For example, a supplier that historically has categorized a given buyer as a friend but subsequently becomes subjected to a stringent monitoring system may reclassify the buyer as a businessperson.[6] It is conceivable that the supplier itself, as a result of role stickiness, will continue to enact the friend role. However, going forward, the supplier's behaviors are likely to be in the form of perfunctory compliance with the buyer's strategy (Moschandreas 1997). For example, the supplier may continue to deliver components according to a fixed delivery schedule but refuse to accommodate special requests during unusual demand conditions (Wathne and Heide 2000). In this scenario, the buyer's governance efforts are actually dysfunctional and associated with opportunity costs because costly initiatives are deployed toward a party that has already internalized a friend role.

Perhaps more seriously, it is also possible that the buyer's use of market-based governance mechanisms in such situations produces actual role switching. The identity gap may lead to a recategorization for the supplier, resulting in a matching businessperson role with an entirely new behavioral repertoire. As a specific example of role switching, consider 7-Eleven's decision to impose an electronic monitoring system on its franchisees. As various sources (e.g., Coughlan et al. 2001; Lal and Han 2005) describe, 7-Eleven franchisees are frequently attracted to the system in the first place because of perceived compatibility between the company's operating philosophy and their own mind-sets. However, the implementation of a "draconian" electronic surveillance system was viewed as such a fundamental shift in management philosophy that it caused severe reactions among the franchisees, some of which were subsequently described as "militant radicals" (Shirouzu and Bigness 1997).

In addition to the direct effects on roles themselves, the mismatch scenarios may produce various long-term outcomes. For example, the new behavioral repertoire that follows from a role switch may include retaliatory actions, which are "characteristically opportunistic" in nature (John 1984, p. 280). Over time, this may have economic consequences in the form of high direct transaction costs in the ongoing relationship.

Importantly, both the general possibility and consequences of role switching are well documented. Kreps (1997, p. 363) notes that a firm that historically has not monitored closely in a given relationship and that has not relied on strong extrinsic incentives may create "muddiness" and a shift in relationship orientation. Tenbrunsel and Messick (1999) show empirically that surveillance and sanctioning systems can change both (1) a party's orientation toward a partner and (2) the actual choice rule used. Specifically, they show that the use of such hard relationship features increased the likelihood that a business type of frame was evoked, which in turn induced a calculative decision rule. Notably, Tenbrunsel and Messick also show that after a business frame was evoked, a weak surveillance and sanctioning system that was intended to increase cooperation actually decreased the likelihood of cooperative behaviors.

The preceding discussion implies that a perceived role switch by an exchange partner may affect a party's own decision logic and that the shift may have nonintuitive (and negative) consequences. In Nagin and colleagues' (2002) terminology, these consequences follow from applying a "rational cheater" model and its underlying assumption (i.e., logic of consequences) to situations in which decision makers are currently working under a "conscience model" or under a logic of appropriateness. Notably, and somewhat disturbingly, Anderson and Jap (2005) cite fieldwork and psychological research that document that close interfirm relationships are particularly vulnerable to showing a "dark side."

---

[6] Technically, such determinations are made on the basis of stylized "if–then" statements (Hayek 1945). For example, if a buyer monitors, then he or she will be recognized as a businessperson.

Existing businessperson role. We now consider a situation in which a party already makes decisions as part of a businessperson role and in which the focal behavioral repertoire is based on a logic of consequences. In Figure 1, the possible scenarios appear in Cells 4, 5, and 6. The matches are represented by Cells 5 and 6, in which incentives and monitoring are deployed toward a party whose current role is consistent with the use of such mechanisms. We expect that the outcome of these scenarios is reinforcement of the current role. Notably, there is evidence that suggests that such matches have distinct performance implications. For example, Tenbrunsel and Messick (1999) show that when a business type of decision frame was evoked in a relationship, strong sanctions indeed increased cooperation.

The mismatch scenario for an existing businessperson role appears in Cell 4. A party whose current mental model is based on a logic of consequences is subsequently subjected to socialization. Although this scenario involves a mismatch, we do not necessarily predict a role switch. Instead, relying on socialization in such a situation may simply cause a resource waste, to the extent that the relevant efforts fail to convert the businessperson to a friend. The magnitude of the socialization efforts needed to induce such a switch is likely to be substantial, as evidenced by Mitsubishi Motor Corporation's efforts to reverse its adversarial supplier relationships (Smitka 2001). To regain its suppliers' support, Mitsubishi Motor Corporation went so far as to establish a formal supplier association and a technical support center staffed with 70 full-time engineers.

Notably, to the extent that a firm's relationship efforts involve dedicated investments, as in the preceding example, a failure to bring about a role switch in a businessperson may produce a mismatch scenario with substantial economic implications. To illustrate, consider Jackson's (1985) analysis of the risks associated with using "relationship marketing" initiatives toward customers that operate under a "transaction" orientation. For example, a supplier that develops a dedicated product and delivery system for a buyer that is looking for only a temporary source of capacity may not only fail to reap the returns from the investments but also subsequently face a significant lock-in condition (Klein, Crawford, and Alchian 1978). This suggests that explicit attention to roles can be an important prerequisite for relationship investment decisions.

Discussion

Theoretical Implications and Relationships with Prior Research

Much of the extant research on interfirm relationships has focused, explicitly or implicitly, on the orientations that parties bring to bear on their interaction. In some of this literature, a sharp distinction has been drawn between different types of orientations. For example, one particular body of research emphasizes calculative orientations, which correspond to March's (1994) logic of consequences. Many of the theoretical frameworks used to study interfirm relationships (e.g., Axelrod 1984; Eisenhardt 1989; Williamson 1985) belong to this category. At the same time, there are theoretical accounts and empirical evidence, primarily from sociology and contract law (e.g., Macaulay 1963; Uzzi 1997) of heuristic orientations, which are consistent with a logic of appropriateness. Overall, however, the extant literature tends to take an "either–or" type of approach, implicitly assuming that relationships can be described in terms of a single orientation.

Our central argument is that such categorical distinctions may be too simplistic. On the one hand, we subscribe to the general distinction between the logics of consequences and appropriateness, as they pertain to individual relationship episodes. As such, our framework is consistent with prior work that describes relationships in terms of locations on particular continuums, such as the degree of trust, commitment, and long-term orientation. Specifically, a high "score" on such a dimension in a given situation would most likely reflect the strength of a prevailing friend role. On the other hand, the upshot of our framework is that descriptions along single dimensions may not fully capture the true nature of relationship interactions. We posit that exchange relationships may simultaneously involve different *types* of orientations. Importantly, this goes beyond acknowledging that relationships are multidimensional, in the sense of joint descriptions on intercorrelated attributes (e.g., different relational norms). Although we endorse this view as it applies to individual transactions, our framework is based on the assumption that actual shifts among fundamentally different relationship orientations are both possible and likely.

Theoretically, our framework begins to provide an interactive perspective on relationship formation and management. For example, a firm may rely simultaneously on (direct) selection processes and on (indirect) opportunities for self-selection through various incentive and monitoring devices. The notion of self-selection suggests that though a particular governance process can be initiated by one party, its ultimate success depends on the other party's tendency to seek out a salient role. As such, our framework suggests that certain governance mechanisms, which have often been described primarily from the perspective of a single firm (i.e., the principal), actually involve and, in a broader sense, are available to both relationship parties. Moreover, interactivity follows from the notion that firms judge a partner's role on the basis of the particular governance choices made, and these judgments ultimately influence the focal firm's own role and relationship decisions. Our framework conceptualizes relationship roles as the products of complex processes that involve multiple parties, governance repertoires, and judgments.

In a recent article, Stout and Blair (2001) criticize certain branches of extant social science research for treating decision making as a "black box" and for being limited to observing particular antecedents and outcomes without accounting for the underlying process. Similar critiques have been voiced by others, who have asked for greater attention to what lies behind choices (Hodgson 1997) and for "mechanism-based explanations" (Hedström and Swedberg 1998, p. 9). In a broad sense, our framework is an attempt to shed some light on this issue. We propose that the concept of a role may yield some insight into the black box and that it may help integrate discrepant theoretical perspectives on relationships and interactions.

Finally, we believe that our role framework has the potential to generate new insights into the unit-of-analysis issue in relationship research. Consider the literature on transaction costs as an example. In this literature, the individual transaction has been suggested as the key unit of analysis (Williamson 1996). Rindfleisch and Heide (1997) note how this assumption has manifested itself in empirical work in the form of a focus on dyadic relationships. Recently, however, this particular focus has been described as being too restrictive, insofar as it ignores the effects of related relationship dyads. As a consequence, researchers have begun to adopt larger units of analysis, such as multiple dyads and networks (e.g., Wathne and Heide 2004).

For all practical purposes, adopting a network perspective means broadening the unit of analysis and taking a macro perspective on relationship research. Although we acknowledge the importance of such a perspective, our framework represents a call for disaggregation, that is, for a micro-level account of relationships and parties. Specifically, our role framework suggests that relationships cannot be properly understood without explicit recognition of both (1) the set of roles with which firms approach relationships in the first place and (2) the ability of relationship management strategies to reinforce or dilute those roles or to cause actual switching. In certain respects, our framework recognizes the "decision premise" as an important unit of analysis, consistent with Simon's (1947) early recommendation.

Importantly, this does not diminish the need to study larger marketing networks. However, accounting for roles suggests that certain phenomena are more complex than frequently assumed. For example, it is possible that the use of particular strategies in one relationship influences not only the pattern of roles in that particular relationship but also the possibility of role switching in related ones. Heide (2003) recently showed that market relationships that coexist with an integrated operation within a so-called plural system exhibit more hierarchical characteristics than singular ones, as a result of spillover effects across connected relationships. In general, we believe that a role-theoretic model of relationship governance may be useful in documenting processes at both a micro and a macro level and both within and across parties and relationships.

Managerial Implications

A considerable body of literature has emerged in marketing on the general topic of governance and relationship management strategies. Both academic research and industry observation suggest that firms' governance decisions are important parts of their overall marketing strategies (Ghosh and John 1999). Moreover, firms' governance choices have been shown to have distinct performance implications (e.g., Buvik and John 2000).

However, the focus in prior research has often been strictly on the governance mechanisms themselves. For example, Baker and Faulkner (1991) develop a set of general recommendations for relationship management, which include the explicit use of incentives ("link pay with performance"), monitoring ("increase accountability"), and related initiatives. Although we do not dispute the possible usefulness of such efforts, we suggest a caveat about the generic use of any governance program. Our conceptual framework accounts for the notion that inherent linkages exist between roles and governance mechanisms and that firms' governance decisions must be made in a context-sensitive manner and with explicit attention to the ability of governance mechanisms both to create roles initially and to activate them over time. In general, understanding relationship roles is an important prerequisite for governance deployment.

In managing relationships, firms must also carefully consider both roles and functions. Consistent with recent role theory, we draw an explicit distinction between the two. For example, in a supply chain context, there is a crucial difference between functions such as inventory maintenance and forecasting and the specific manner in which they are carried out (i.e., as a friend or businessperson). As Kahn (2003) notes, the implementation of efficient-consumer-response and quick-response types of initiatives requires fundamental changes in relationships between supply chain partners with respect to both the allocation of functions and overall relationship orientation.

To some extent, the focus of our conceptual framework is on mismatches between governance decisions and roles and on their potential outcomes. For example, we suggest that incentives and monitoring, which have been recommended as generic relationship management initiatives (e.g., Baker and Faulkner 1991), are intrinsically linked with only one type of role, namely, that of a businessperson. Crucially, deploying such strategies in other contexts (i.e., toward friends) not only may be ineffective or cause resource waste but also may cause role switching and subsequent activation of behavioral repertoires that may include opportunism.

Notably, although the concept of relationship roles has received little systematic attention in extant research, it is frequently suggested as a key explanation of relationship problems. For example, Solomon and colleagues (1985, p. 105) note how a "failure to read from a common script" is a typical cause of difficulties in relationships between clients and service providers. Similarly, Uzzi (1999, p. 502) concludes that it is inherently difficult to form ties between parties that "use alternate scripts." Although more fine-grained guidelines for managerial decision making require a body of empirical evidence, our general conclusion is that firms' governance decisions are inherently more complex than is frequently assumed and that such decisions may hinge crucially on detailed assessments of relationship roles.

Issues for Further Research

Our conceptual framework in Figure 1 involves a series of research issues and hypotheses that can be tested empirically. In general, the underlying philosophy of the framework is that linkages exist among (1) *ex ante* roles or "initial conditions" (Cowen 1998), (2) governance strategies, and (3) outcomes. Clarification of the specific nature of these linkages requires empirical testing, including efforts to measure directly and account for relationship roles. We echo Kreps's (1997, p. 364) sentiments that such empirical efforts are "likely to be messy." Nonetheless, increasing the understanding of roles and their interplay with established relationship constructs seems to be an important research priority.

In addition to the general research agenda inherent in our conceptual framework, there are additional, more specific research issues that should be systematically examined. We discuss some of them next.

Relationship investments and resource allocation decisions. Important issues pertain to the linkage between roles and profitability. Most likely, the two particular roles we discuss herein differ substantially in terms of their cash-flow patterns. For example, creating a friend role may require significant up-front investments, for example, in the form of selection and socialization efforts. At the same time, after a friend is established, it may be cheaper to maintain than a businessperson. However, note that certain types of exchange partners that may fit a friend profile can become demanding and costly to serve (Reinartz and Kumar 2003).

It is also important to document how strongly roles are held and how easily they can be switched or "unfrozen." For example, if certain customers have undesirable cost implications, as suggested in Reinartz and Kumar's (2003) research, and if established roles are indeed sticky, it suggests that a firm's role creation decisions must be approached in a deliberate and forward-looking manner.

At the same time, although existing sticky roles may have nontrivial cost implications, they may also represent sources of competitive advantage. For example, consider an industrial supplier that enjoys sticky friend relationships with its buyers. From the standpoint of a competing supplier, this raises important questions about both (1) the specific sources of role stickiness and (2) the particular "corrective procedures" (Bolton and Reed 2004) that may undermine the incumbent supplier's relationships.

Finally, a related profitability issue pertains to resource allocation decisions across relationship types within an overall portfolio (Johnson, Sohi, and Grewal 2004; Johnson and Selnes 2004). At a given time, a firm may possess a mix of roles in its portfolio, and decisions must be made about how to allocate scarce resources among them. Dickson and colleagues (2004) note how firms may need to make fundamental allocation choices between (1) using scarce resources to reinforce prior behavior (e.g., in a current friend relationship) and (2) using them to solve problems in others (e.g., trying to convert a businessperson). These types of issues require firms to consider both initiation and maintenance costs carefully, at the relationship and the portfolio levels.

Relationship perceptions. As we noted previously, deploying governance mechanisms without a careful consideration of existing relationship roles may have unintended consequences, such as resource waste and role switching. Given the well-documented finding that firms have divergent perceptions about their interaction (e.g., Wathne, Biong, and Heide 2001), such scenarios are not unlikely. Indeed, on the basis of previous research, it might be expected that the potential for mismatches between governance deployment and prevailing roles is high and that this may be a possible explanation of relationship failure. Research could be directed usefully toward documenting (1) the manner in which role perceptions are established and (2) whether particular roles are uniquely held by an individual party or are shared in some way. A related question is whether a party may misrepresent a given role. In *The Republic*, Plato discusses how Socrates defines a friend as "one who both seems and is an honest man."[7] It is clear that distinct opportunities exist, at least in the short run, for "seeming" and for the misrepresentation of roles. This raises important relationship management challenges in its own right.

In general, role theory suggests that a distinction can be made between perceptions that pertain to (1) a specific exchange partner, as per our framework, and (2) certain positions or organizational forms (Biddle 1979; Forman and Whetten 2002), as per conventional role theory. For example, a manufacturer may have a perception of retailers in general, which will influence its approach to a particular partner. As a specific example, a manufacturer whose prior interaction has primarily involved smaller retail clients may perceive retailers in general as friends and therefore approach new retail relationships with this particular orientation. Our framework highlights the risks associated with such mental shortcuts, insofar as they produce costly mismatched governance initiatives.

Issues such as those we present in the preceding discussion illustrate how roles can be conceptualized at multiple levels of analysis. With increasing demand for cross-level constructs and theories (e.g., Klein and Kozlowski 2000; Rousseau 1985), role theory represents a promising avenue for further research. Our discussion also reinforces the importance of continuing to direct relationship research toward improving the understanding of human behavior. Notably, Williamson (1996, p. 266), who is frequently criticized for the particular behavioral assumptions that underlie transaction cost economics, explicitly suggests that "more microanalytic attention to the processes through which trading relationships evolve is indeed a rewarding research enterprise." Similarly, Tenbrunsel and Messick (1999, p. 706) emphasize the "need to explore assumptions" about decision makers and their reactions to the mechanisms that are deployed in exchange relationships. Although continued theoretical integration and empirical research may turn out to be messy, the resultant insights could be considerable.

---

[7] See Plato (2000), *The Republic*, trans. Tom Griffith. Cambridge, UK: Cambridge University Press.

REFERENCES

Albert, Stuart and David A. Whetten (1985), "Organizational Identity," in *Research in Organizational Behavior*, Vol. 7, Larry L. Cummings and Barry M. Staw, eds. Greenwich, CT: JAI Press, 263–95.

Anderson, Erin and Sandy D. Jap (2005), "The Dark Side of Close Relationships," *Sloan Management Review*, 46 (3), 75–82.

——— and Richard L. Oliver (1987), "Perspectives on Behavior-Based Versus Outcome-Based Salesforce Control Systems," *Journal of Marketing*, 51 (October), 76–88.

Anderson, James C., Håkan Håkansson, and Jan Johanson (1994), "Dyadic Business Relationships Within a Business Network Context," *Journal of Marketing*, 58 (October), 1–15.

Andreoni, James (1989), "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," *Journal of Political Economy*, 97 (6), 1447–58.

Aron, Laurie Joan (1998), "More Carrot, Less Stick for Vendors," *Chain Store Age*, 74 (12), 232–34.

Ashforth, Blake E. (2001), *Role Transitions in Organizational Life: An Identity-Based Perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.

Axelrod, Robert (1984), *The Evolution of Cooperation*. New York: Basic Books.

Baker, Wayne A. (1990), "Market Networks and Corporate Behavior," *American Journal of Sociology*, 96 (3), 589–625.

——— and Robert R. Faulkner (1991), "Strategies for Managing Suppliers of Professional Services," *California Management Review*, 33 (4), 33–45.

Barney, Jay B. and Mark H. Hansen (1994), "Trustworthiness as a Source of Competitive Advantage," *Strategic Management Journal*, 15 (Special Issue), 175–90.

Bates, Frederick L. and Clyde C. Harvey (1975), *The Structure of Social Systems*. New York: Gardner Press.

Becker, Gary S. (1976), *An Economic Approach to Human Behavior*. Chicago: University of Chicago Press.

Behrman, Douglas N. and William D. Perreault Jr. (1984), "A Role Stress Model of the Performance and Satisfaction of Industrial Salespersons," *Journal of Marketing*, 48 (Fall), 9–21.

Biddle, Bruce J. (1979), *Role Theory: Expectations, Identities, and Behaviors*. New York: Academic Press.

——— (1986), "Recent Developments in Role Theory," *Annual Review of Sociology*, 12, 37–59.

Blau, Judith R. (1993), *Social Contracts and Economic Markets*. New York: Plenum Press.

Bolton, Lisa E. and Americus Reed II (2004), "Sticky Priors: The Perseverance of Identity Effects on Judgments," *Journal of Marketing Research*, 41 (October), 397–410.

Brickley, James A. and Frederick H. Dark (1987), "The Choice of Organizational Form: The Case of Franchising," *Journal of Financial Economics*, 18 (2), 401–420.

Brooks, Rick (2004), "A Deal with Target Put Lid on Revival at Tupperware," *The Wall Street Journal*, (February 18), A1.

Burt, Ronald S. (1982), *Toward a Structural Theory of Action: Network Models of Social Structure, Perception, and Action*. New York: Academic Press.

Buvik, Arnt and George John (2000), "When Does Vertical Coordination Improve Industrial Purchasing Relationships," *Journal of Marketing*, 64 (October), 52–64.

Camerer, Colin and Ari Vepsalainen (1988), "The Economic Efficiency of Corporate Culture," *Strategic Management Journal*, 9 (Special Issue), 115–26.

Cannon, Joseph P. and William D. Perreault Jr. (1999), "Buyer–Seller Relationships in Business Markets," *Journal of Marketing Research*, 36 (November), 439–60.

Chatman, Jennifer A. (1991), "Matching People and Organization: Selection and Socialization in Accounting Firms," *Administrative Science Quarterly*, 36 (September), 459–94.

Churchill, Gilbert A., Neil M. Ford, Steven W. Hartley, and Orville C. Walker (1985), "The Determinants of Salesperson Performance: A Meta-Analysis," *Journal of Marketing Research*, 22 (May), 103–118.

Corporate University Review (2000), "Six Sigma Institute at Motorola University," 8 (3), 12–14.

Coughlan, Anne T., Erin Anderson, Louis W. Stern, and Adel I. El-Ansary (2001), *Marketing Channels*. Upper Saddle River, NJ: Prentice Hall.

Cowen, Tyler (1998), "Do Economists Use Social Mechanisms to Explain?" in *Social Mechanisms: An Analytical Approach to Social Theory*, Peter Hedström and Richard Swedberg, eds. New York: Cambridge University Press, 125–46.

Dickson, Peter R., Nancy Buchan, Diana L. Haytko, and Jan B. Heide (2004), "Subjective Probability and the 'Is' and 'Ought' of Relationship Investments," working paper, Department of Marketing, Florida International University.

Dubinsky, Alan J., Roy D. Howell, Thomas N. Ingram, and Danny N. Bellenger (1986), "Sales Force Socialization," *Journal of Marketing*, 50 (October), 192–207

Durkheim, Emilie (1915), *The Elementary Forms of the Religious Life*. New York: The Free Press

Dutta, Shantanu, Mark Bergen, and George John (1994), "The Governance of Exclusive Territories When Dealers Can Bootleg," *Marketing Science*, 13 (1), 83–99.

Dwyer, F. Robert, Paul H. Schurr, and Sejo Oh (1987), "Developing Buyer–Seller Relationships," *Journal of Marketing*, 51 (April), 11–27.

The Economist (1999), "The Burger King," (October), 78.

Eisenhardt, Kathleen M. (1989), "Agency Theory: An Assessment and Review," *Academy of Management Review*, 14 (1), 57–74.

El-Ansary, Adel I. and Louis W. Stern (1972), "Power Measurement in the Distribution Channel," *Journal of Marketing Research*, 9 (February), 47–52.

Elster, John (1989), "Social Norms and Economic Theory," *Journal of Economic Perspectives*, 3 (4), 99–117.

Etzioni, Amitai (1975), *A Comparative Analysis of Complex Organizations*. New York: MacMillan.

Forgas, Joseph P. (1982), "Episode Cognition: Internal Representations of Interaction Routines," in *Advances in Experimental Social Psychology*, Leonard Berkowitz, ed. New York: Academic Press, 59–103.

Forman, Peter and David A. Whetten (2002), "Members' Identification with Multiple-Identity Organizations," *Organization Science*, 13 (6), 618–35.

Frazier, Gary L. (1999), "Organizing and Managing Channels of Distribution," *Journal of the Academy of Marketing Science*, 27 (2), 226–41.

Friedberg, Erhard (2000), "Going Beyond the Either/Or," *Journal of Management and Governance*, 4 (1–2), 35–52.

Gambetta, Diego (1996), *Were They Pushed or Did They Jump? Individual Decision Mechanisms in Education*. New York: Cambridge University Press.

Gardiner, Morse (2003), "Why We Misread Motives," *Harvard Business Review*, 81 (1), 18.

Ghosh, Mrinal and George John (1999), "Governance Value Analysis and Marketing Strategy," *Journal of Marketing*, 63 (Special Issue), 131–45.

Gibbons, Robert (1999), "Taking Coase Seriously," *Administrative Science Quarterly*, 44 (1), 145–57.

Gilliland, David I. (2003), "Toward a Business-to-Business Channel Incentives Classification Scheme," *Industrial Marketing Management*, 32 (1), 55–67.

Goffman, Erving (1969), *Strategic Interaction*. Philadelphia: University of Pennsylvania Press.

Granovetter, Mark (1985), "Economic Action and Social Structure: The Problem of Embeddedness," *American Journal of Sociology*, 91 (November), 48–510.

Greenwald, Anthony G. and Anthony R. Pratkanis (1984), "The Self," in *Handbook of Social Cognition*, Robert S. Wyer and Thomas K. Srull, eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 129–78.

Hackett, Steven C. (1994), "Is Relational Exchange Possible in the Absence of Reputations and Repeated Contract?" *Journal of Law, Economics, and Organization*, 10 (2), 360–89.

Harrison, Richard J. and Glenn R. Caroll (1991), "Keeping the Faith: A Model of Cultural Transmission in Formal Organizations," *Administrative Science Quarterly*, 36 (4), 552–82.

Harsanyi, John C. (1955), "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, 63 (August), 309–321.

Hayek, Friedrich (1945), "The Use of Knowledge in Society," *American Economic Review*, 35 (4), 519–30.

Hedström, Peter and Richard Swedberg (1998), *Social Mechanisms: An Analytical Approach to Social Theory*. New York: Cambridge University Press.

Heide, Jan B. (1994), "Interorganizational Governance in Marketing Channels," *Journal of Marketing*, 58 (January), 71–85.

——— (2003), "Plural Governance in Industrial Purchasing," *Journal of Marketing*, 67 (October), 18–29.

——— and George John (1992), "Do Norms Matter in Marketing Relationships?" *Journal of Marketing*, 56 (April), 32–44.

Hodgson, Geoffrey M (1997), "The Ubiquity of Habits and Rules," *Cambridge Journal of Economics*, 21 (6), 663–84.

Hogarth, Robin M. and Melvin W. Reder (1987), *Rational Choice: The Contrast Between Economics and Psychology*. Chicago: University of Chicago Press.

Iacobucci, Dawn (1996), *Networks in Marketing*. Thousand Oaks, CA: Sage Publications.

Jackson, Barbara B. (1985), *Winning and Keeping Industrial Customers*. Lexington, MA: Lexington Books.

Jensen, William H. and Michael C. Meckling (1976), "Theory of the Firm· Managerial Behavior, Agency Costs, and Ownership Structure," *Journal of Financial Economics*, 3 (October), 305–360.

John, George (1984), "An Empirical Investigation of Some Antecedents of Opportunism in a Marketing Channel," *Journal of Marketing Research*, 21 (August), 278–89.

——— and Barton Weitz (1989), "Sales Force Compensation: An Empirical Investigation of Factors Related to Use of Salary Versus Incentive Compensation," *Journal of Marketing Research*, 26 (February), 1–14.

Johnson, Jean L., Ravipreet S. Sohi, and Rajdeep Grewal (2004), "The Role of Relational Knowledge Stores in Interfirm Partnering," *Journal of Marketing*, 68 (July), 21–36.

Johnson, Michael D. and Fred Selnes (2004), "Customer Portfolio Management: Toward a Dynamic Theory of Exchange Relationships," *Journal of Marketing*, 68 (April), 1–17.

Kahn, Gabriel (2003), "Invisible Supplier Has Penney's Shirts All Buttoned Up," *The Wall Street Journal*, (September 11), A1, A9.

Kahn, Robert L., Donald M. Wolfe, Robert P. Quinn, J. Diedrick Snoek, and Robert A. Rosenthal (1964), *Organizational Stress: Studies in Role Conflict and Ambiguity*. New York: John Wiley & Sons.

Killian, Erin (2003), "Butter' Em Up," *Forbes*, (June 9), 175–76.

Klein, Benjamin (1996), "Why Hold-Ups Occur: The Self-Enforcing Range of Contractual Relationships," *Economic Inquiry*, 34 (3), 444–63.

———, Robert A. Crawford, and Arman A. Alchian (1978), "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," *Journal of Law and Economics*, 21 (October), 297–326.

Klein, Katherine J. and Steve W. Kozlowski (2000), *Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions*. San Francisco: Jossey-Bass.

Kohli, Ajay K. and Bernard J. Jaworski (1990), "Market Orientation: The Construct, Research Propositions, and Management Implications," *Journal of Marketing*, 54 (April), 1–18.

Kreps, David M. (1997), "The Interaction Between Norms and Economic Incentives: Intrinsic Motivation and Extrinsic Incentives," *American Economic Review*, 87 (2), 359–64.

Lal, Rajiv and Arar Han (2005), "Tanpin Kanri: Retail Practice at Seven-Eleven Japan," Harvard Business School Case No. 9-506-002.

Larson, Andrea (1992), "Network Dyads in Entrepreneurial Settings: A Study of the Governance of Exchange Relationships," *Administrative Science Quarterly*, 37 (1), 76–104.

Lyons, Thomas F., Richard A. Krachenberg, and John W. Henke (1990), "Mixed Motive Marriages: What's Next for Buyer-Supplier Relationships?" *Sloan Management Journal*, 31 (3), 29–36.

Macaulay, Stewart (1963), "Non-Contractual Relations in Business: A Preliminary Study," *American Sociological Review*, 28 (February), 55–69.

March, James G. (1994), *A Primer on Decision Making: How Decisions Happen*. New York: The Free Press.

——— (1999), *The Pursuit of Organizational Intelligence*. Malden, MA: Blackwell.

Margolis, Howard (1982), *Selfishness, Altruism, and Rationality: A Theory of Social Choice*. Chicago: University of Chicago Press

Margonelli, Lisa (2002), "How Ikea Designs Its Sexy Price Tags," *Business 2.0*, (October), 1–7.

Mead, George M. (1932), *Mind, Self & Society from the Standpoint of a Social Behaviorist*. Chicago: University of Chicago Press.

Meister, Jeanne C. (1998), *Corporate Quality Universities: Lessons in Building a World-Class Work Force*. New York: McGraw-Hill.

Messick, David M. (1999), "Alternative Logics for Decision Making in Social Settings," *Journal of Economic Behavior & Organization*, 39 (1), 11–28.

Montgomery, James D. (1998), "Toward a Role-Theoretic Conception of Embeddedness," *American Journal of Sociology*, 104 (1), 92–125.

Moschandreas, Maria (1997), "The Role of Opportunism in Transaction Cost Economics," *Journal of Economic Issues*, 31 (1), 39–57.

Nagin, Daniel S., James B. Rebitzer, Seth Sanders, and Lowell J. Taylor (2002), "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment," *American Economic Review*, 92 (4), 850–73.

Normann, Richard and Rafael Ramirez (1993), "From Value Chain to Value Constellation: Designing Interactive Strategy," *Harvard Business Review*, 71 (4), 65–77.

Oakes, Penelope J. (1987), "The Salience of Social Categories," in *Rediscovering the Social Group*, John C. Turner, ed. New York: Basil Blackwell, 117–41.

Ouchi, William G. (1979), "A Conceptual Framework for the Design of Organizational Control Mechanisms," *Management Science*, 25 (9), 833–48.

Parsons, Talcott (1937), *The Structure of Social Action*. New York: McGraw-Hill.

Powell, Walter W. and Laurel Smith-Doerr (1994), "Networks and Economic Life," in *Business Marketing: An Interaction and Network Perspective*, Neil Smelser and Richard Swedberg, eds. Princeton, NJ: Princeton University Press, 368–402.

Purchasing (1995), "Big Three Name Quality Registrars," (April 6), 57.

Reinartz, Werner and V. Kumar (2003), "The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration," *Journal of Marketing*, 67 (January), 77–99.

Ridgway, Valentine F. (1957), "Administration of Manufacturer-Dealer Systems," *Administrative Science Quarterly*, 1 (March), 464–83.

Rindfleisch, Aric and Jan B. Heide (1997), "Transaction Cost Analysis: Past, Present, and Future Applications," *Journal of Marketing*, 61 (October), 30–54.

Rokkan, Aksel I, Jan B. Heide, and Kenneth H. Wathne (2003), "Specific Investments in Marketing Relationships: Expropriation and Bonding Effects," *Journal of Marketing Research*, 40 (May), 210–24.

Rousseau, Denise M. (1985), "Issues of Level in Organizational Research: Multi-Level and Cross-Level Perspectives," in *Research in Organizational Behavior*, Barry M. Staw and Larry L. Cummings, eds. Greenwich, CT: JAI Press, 1–37.

Ryan, Richard M. and Edward L. Deci (2000), "Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions," *Contemporary Educational Psychology*, 25 (1), 54–67.

Salancik, Gerald R. (1995), "Wanted: A Good Network Theory of Organization," *Administrative Science Quarterly*, 40 (June), 345–49.

Schick, Shane (2001), "Cisco Takes Stock of Resellers," *Computer Dealer News*, (August 31), 34.

Scott, John F. (2000), "Rational Choice Theory," in *Understanding Contemporary Society: Theories of the Present*, Abigail Halcli, Gary Browning, and Frank Webster, eds. London: Sage Publications, 126–38.

Sherif, Muzafer (1936), *The Psychology of Social Norms*. New York: Harper.

Shirouzu, Norihiko and Jon Bigness (1997), "7-Eleven Operators Resist System to Monitor Managers," *The Wall Street Journal*, (June 16), B1, B5.

Simon, Herbert A. (1947), *Administrative Behavior*. New York: The Free Press.

Singh, Jagdip (2000), "Performance Productivity and Quality in Frontline Employees in Service Organizations," *Journal of Marketing*, 64 (April), 15–34.

Smitka, Michael J. (2001), *Competitive Ties: Subcontracting in the Japanese Automotive Industry* New York: Columbia University Press.

Solomon, Michael R., Carol Surprenant, John A. Czepiel, and Evelyn G. Gutman (1985), "A Role Theory Perspective on Dyadic Interactions: The Service Encounter," *Journal of Marketing*, 49 (Winter), 99–111.

Stern, Louis W. (1969), *Distribution Channels: Behavioral Dimensions*. Boston: Houghton Mifflin.

Stout, Lynn A. and Margaret M. Blair (2001), "Trust, Trustworthiness, and the Behavioral Foundations of Corporate Law," *University of Pennsylvania Law Review*, 149 (6), 1735–1810.

Stryker, Sheldon (1968), "Identity Salience and Role Performance: The Relevance of Symbolic Interaction Theory for Family Research," *Journal of Marriage and the Family*, 4 (30), 558–64.

Tajfel, Henri and John C. Turner (1986), "The Social Identity Theory of Intergroup Behavior," in *Advances in Group Processes*, Stephen Worchel and William G. Austin, eds. Chicago. Nelson-Hall, 77–122.

Tatge, Mark (2001), "McMissteps," *Forbes*, (December 10), 76–78.

Telser, Lester G. (1980), "A Theory of Self-Enforcing Agreements," *Journal of Business*, 53 (1), 27–44.

Tenbrunsel, Ann E. and David M. Messick (1999), "Sanctioning Systems, Decision Frames, and Cooperation," *Administrative Science Quarterly*, 44 (4), 684–707.

Turner, John C., Michael A. Hogg, Penelope J Oakes, Stephen D. Reicher, and Margaret S. Wetherell (1987), *Rediscovering the Social Group: A Self-Categorization Theory*. New York: Basil Blackwell.

Turner, Jonathan H. (1974), *The Structure of Sociological Theory*. Chicago: University of Chicago Press.

Uzzi, Brian (1997), "Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness," *Administrative Science Quarterly*, 42 (March), 35–67.

——— (1999), "Embeddedness in the Making of Financial Capital: How Social Relations and Networks Benefit Firms Seeking Financing," *American Sociological Review*, 64 (4), 481–505.

Wathne, Kenneth H., Harald Biong, and Jan B. Heide (2001), "Supplier Choice in Embedded Markets: Relationship and Marketing Program Effects," *Journal of Marketing*, 65 (April), 54–66.

——— and Jan B. Heide (2000), "Opportunism in Interfirm Relationships: Forms, Outcomes, and Solutions," *Journal of Marketing*, 64 (October), 36–51.

——— and ——— (2004), "Relationship Governance in a Supply Chain Network," *Journal of Marketing*, 68 (January), 73–89.

Webster, Fredrick E. and Yoram Wind (1972), *Organizational Buying Behavior*. Englewood Cliffs, NJ: Prentice Hall.

Williamson, Oliver E. (1975), *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: The Free Press.

——— (1985), *The Economic Institutions of Capitalism*. New York: The Free Press.

——— (1993), "Calculativeness, Trust, and Economic Organization," *Journal of Law and Economics*, 36 (1), 453–86.

——— (1996), *The Mechanisms of Governance*. New York: Oxford University Press.

Winship, Christopher and Michael Mandel (1983), "Roles and Positions: A Critique and Extension of the Blockmodeling Approach," in *Sociological Methodology, 1983–1984*, Samuel Leinhardt, ed. San Francisco: Jossey-Bass, 314–44

Matthew Thomson

Human Brands: Investigating Antecedents to Consumers' Strong Attachments to Celebrities

This article explores recent advances in self-determination research to address why consumers develop strong attachments to "human brands," a term that refers to any well-known persona who is the subject of marketing communications efforts. Study 1 uses a survey that is analyzed with structural equation modeling. Study 2 is qualitative and offers corroborating evidence for the proposed theoretical model. Study 3 extends the model with a more naturalistic sample and tests several alternative hypotheses using hierarchical regression. The results suggest that when a human brand enhances a person's feelings of autonomy and relatedness and does not suppress feelings of competence, the person is likely to become more strongly attached to it. This article documents that strong attachments are predictive of satisfied, trusting, and committed relationships and proposes that attachment strength may be a parsimonious proxy for consumer–brand relationship strength. The results imply that benefits would accrue to organizations such as entertainment firms and political parties that establish direct and routine interaction between human brands and consumers, that human brands to which consumers are attached offer significant potential as endorsers, and that organizations should address how to make the human brands they manage more authentic.

> "When I go to the grocery store, people hug me. But that's OK. People do know me. They've seen me grow up on television."
>
> —Michelle Kwan[1]

In the past decade, academic and practitioner interest in consumer relationships has flourished (e.g., Sheth and Parvatiyar 1995). One approach to studying these relationships focuses on feelings of attachment and love (e.g., Carroll and Ahuvia 2006), which lay "at the core of all strong brand relationships" (Fournier 1998, p. 363).

This article focuses on consumers' attachments to human brands, a term that refers to any well-known persona who is the subject of marketing communications efforts (e.g., Rindova, Pollock, and Hayward 2006). The central purpose of this article is to address why consumers form strong attachments to human brands. From a practical standpoint, advancing the understanding of the relationship between consumers and human brands is important because much of the success of the $190 billion U.S. media and entertainment industry hinges on the successful positioning of one of its key assets, the celebrity (Plunkett Research 2004). Organizations spend vast sums annually in an effort to establish psychological connections between consumers and human brands, such as Michael Jordan, Tiger Woods, and Oprah (Serwer 2001). Firms even specialize in brand management with respect to dead celebrities, such as Steve McQueen and John Wayne (Halpern 2005). Similarly, in the 2004 U.S. presidential election, each of the two main political parties spent in excess of $1 billion dollars in an effort to position their candidates favorably with voters (Edsall and Grimaldi 2004). That is, organizations are already dedicated to managing human brands and building emotional bonds with consumers. This article is an effort to explore this phenomenon methodically; it does so by advancing hypotheses and testing them in three studies. Study 1 uses a survey that is analyzed with structural equation modeling. Study 2 is qualitative and offers corroborating evidence for the proposed theoretical model. Study 3 extends the model with a more naturalistic sample and tests several alternative hypotheses using hierarchical regression.

Conceptual Foundation

Human Brands and the Branding Literature

> "Your client, whether they are an athlete or an actor or an actress, has intangible assets: a name, a reputation, a credibility and an image. All of those attributes may be combined into something that could be made into a brand."
>
> —Brian Dubin (quoted in Towle 2003)

This article's underlying premise is that human brands may be viewed as one of several operationalizations of the broader concept of a brand. In marketing, the term "brand" is typically applied to firms, products, and services, and in general, marketers accept that brands may be described in

---

[1] Michelle Kwan, world champion figure skater, quoted in Horovitz (2003).

Matthew Thomson is an assistant professor, Queen's School of Business (e-mail: mthomson@business.queensu.ca). The author thanks Debbie MacInnis, C.W. Park, Brian Lickel, Allen Weiss, Jay Handelman, Peter Dacin, Allison Johnson, and the anonymous *JM* reviewers for feedback on previous drafts of this article.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

terms of perceived quality, image, and so forth. Celebrities can also be considered brands because they can be professionally managed and because they have additional associations and features of a brand. Evidence for this idea can perhaps be most readily found in the context of political campaigns, in which the candidate's message, public appearance, endorsements, and so on, are all controlled by consultants and political parties in the hopes of managing perceived quality and brand image to increase "market share" at the polls (Simon, Gilgoff, and Samuel 2004). Likewise, the National Basketball Association recently established an off-the-court dress code for its players in an effort to improve the league's appeal by managing the image of its individual human brands (Lee 2005). However, many of the players (e.g., Allen Iverson) had previously established and lucrative images that were misaligned with the new dress code; this underscores the importance of understanding how to better manage the brand equity of individual human brands.

Why Attachments Matter

Considerable research has examined features of the various types of relationships that people form with human brands. For example, idolatry (Houran, Navik, and Zerrusen 2005), fandom (Leets, De Becker, and Giles 1995), and celebrity worship (Dietz et al. 1991) have been explored. Much of this research has been advanced under the rubric of attachment theory and its associated literature, which is both well articulated and relevant to marketing (Kleine and Baker 2004). Attachments are a type of strong relationship that people usually first experience as children with their parents; later in life, these attachments routinely develop with other "targets," such as human brands (Leets, De Becker, and Giles 1995). A person immersed in such an emotionally significant relationship normally perceives the relationship partner as differentiated and irreplaceable (i.e.,, target specific; see Hazan and Shaver 1994). When these types of relationships are experienced in reference to human brands, they are typically referred to as "secondary object" attachments and have been described as "intimacy at a distance" (Horton and Wohl 1956).

This article focuses on the strength of attachment, defined as the intensity of a person's target-specific emotional bond with a human brand. Prior research reveals that attachments are differentiated from other constructs. For example, the strength of an attachment is orthogonal to involvement, satisfaction, loyalty, and attitude favorability (Ambler et al. 2002; Thomson, MacInnis, and Park 2005).

This literature suggests that attachment theory can make a contribution to marketing because of the distinctive qualities of an attachment. For example, much as marketers may attempt to create relationships that are trusting, committed, and satisfied (Fournier, Dobscha, and Mick 1998), people routinely report elevated levels of each in their attachments to a variety of objects (Rempel, Ross, and Holmes 2001; Spake et al. 2004). Prior literature points to the independence of attachment strength from trust, satisfaction, and commitment, but the correlation among these constructs suggests that understanding how to create or intensify attachments could offer both an effective and an economical means of achieving stronger marketing relationships. Indeed, Fournier's (1998) article on consumer–brand relationships suggests that feelings linked to attachments are fundamental to strong brand relationships. Therefore, attachment strength may provide a parsimonious and unidimensional indicator of "relationship quality" or strength.

Furthermore, more intense attachments are linked to several consequences that are desirable to marketers. For example, attachments in a marketing context may prevent consumer defections (Liljander and Strandvik 1995), increase consumers' forgiveness in the face of negative information (Ahluwalia, Unnava, and Burnkrant 2001), and predict brand loyalty and willingness to pay (Thomson, MacInnis, and Park 2005). If marketers understand what determines the strength of attachments, they may be in a better position to foster more durable relationships with consumers.

Finally, research implies that attachments may be pivotal to understanding customer-based brand equity, in which certain brands "resonate" with consumers and indicate financial value for the firm (Ambler et al. 2002; Keller 2001). For various reasons, therefore, contemplating what strengthens attachments seems to be an endeavor that is worthwhile from both a practical and a theoretical standpoint.

Relationship similarities. In some respects, relationships involving human brands are a hybrid of other relationships. For example, they are pertinent to marketing because of the central role of the consumer, but they are also a relevant extension of interpersonal research because they implicate a bona fide human being (e.g., David Beckham, Hilary Swank), not a comparatively inanimate object (e.g., Tide laundry detergent). Prior research has considered celebrities, well-known figures, and group entities, such as sports teams or musical acts, and has documented how they are similar to interpersonal relationships in many ways. For example, people often experience "seeming face-to-face" relationships with human brands who are "met as if they were in the circle of one's peers" (Horton and Wohl 1956, p. 215). These relationships evince many of the same expectations, cognitions, emotions, and behaviors that operate in normal interpersonal relationships to the point that a consumer might view a human brand as a pleasant companion, good friend, or romantic mate (Cole and Leets 1999; Rubin and McHugh 1987). Many of the same variables that provide the basis for interpersonal attachments operate with respect to human brands, such as increasing perceptions of a shared background and elevated social appeal (Cole and Leets 1999; Perse and Rubin 1989).

Relationship differences. Although research in marketing has embraced the idea that consumer relationships may be analogous to interpersonal bonds (Fournier 1998), there are differences between the two; however, in general, these important difference are underresearched. For example, Rubin and McHugh (1987) note that unlike interpersonal relationships, people's relationships with celebrities are less likely to be truly interactive (e.g., mutual self-disclosure and interrogation are not possible). However, the effect of these and other differences in relationship functioning is not well understood (Kleine and Baker 2004). Until their impact is addressed, comprehension of how consumers' human-brand attachments form may be advanced by attending to theory development that is dependent on similarities between these contexts. That is, many of the same causes of strong attachments operate when the focal object is a person, a brand, or some hybrid of the two (i.e., human brands).

What Strengthens Attachments?

In general, it is construed that the chief function of attachments is to confer emotional security to the attached party by being responsive to a person's needs (Hazan and Shaver 1994). This theory is expanded on by recent work on fundamental human needs that suggests that if an object is responsive to a person's needs for autonomy, relatedness, and competence (i.e., A-R-C), intense attachments may result (La Guardia et al. 2000; Ryan and Deci 2000).

Defining "autonomy." Autonomy refers to a person's need to feel that his or her activities are self-chosen, self-governed, and self-endorsed (Deci and Ryan 2000; Ryan and Deci 2000). Associated with the fulfillment of this need is a person's perception that he or she is free from pressure to behave in certain ways and is able to express him- or herself as he or she wishes. A person whose need for autonomy is satiated is likely to report feelings of "volition, agency, and initiative" (La Guardia et al. 2000, p. 368). Autonomy is similar to what Austin and Vancouver (1996, p. 357) label "self-determination," that is, a person's sense of freedom to make his or her own choices and to avoid feeling constrained or coerced.

Defining "relatedness." Relatedness refers to a person's need to feel a sense of closeness with others (Deci and Ryan 2000). The need for relatedness is a homonomous tendency; it is a desire to belong to a social sphere (Ryan and Deci 2000) and to avoid feeling isolated (Austin and Vancouver 1996). A person whose need for relatedness is satisfied is likely to report feeling "connected with and cared for by another" (La Guardia et al. 2000, p. 368).

Defining "competence." Competence refers to a person's innate, life-span tendency to seek feelings of effectiveness, achievement, and challenge in his or her activities (Deci and Ryan 2000). This idea is similar to what previous researchers have termed "mastery," which refers to a person's avoidance of mediocrity and decrements in performance (Austin and Vancouver 1996). A person whose need for competence is satiated will report feeling curious and skilled (La Guardia et al. 2000).

The identification of these three variables emerges from recent studies on human motivation, and "although motivation is often treated as a singular construct, even superficial reflection suggests that people are moved to act by very different types of factors" (Ryan and Deci 2000, p. 69). Rather than considering motivation as rising from general human needs, the current research attempts to parcel out the specific motivations and needs that are active in particular social spheres by "asking what kind of motivation is being exhibited at any given time" (Ryan and Deci 2000, p. 69). As such, this article proposes that the consumer–human-brand dyad may describe a relationship context in which these three needs are potentially important because their fulfillment in that context may lead to carefully targeted feelings of attachment.

A-R-C as responsiveness. Autonomy, relatedness, and competence are fundamental human needs (Deci and Ryan 2000). They are different from most needs because they are universal, innate, and enduring. However, this article is not interested in the existence of these needs but rather in the notion that specific partners, such as human brands, might provide an individual with relational inputs to satisfy these needs; this concept is referred to as "responsiveness."

The idea of responsiveness is foundational to attachment research, and self-determination theorists propose that the fulfillment of the A-R-C needs qualifies as responsiveness (e.g., Deci and Ryan 2000). Responsive relational partners "are ones who respond in ways that promote a person's experienced satisfaction of these basic [A-R-C] psychological needs" (La Guardia et al. 2000, p. 368). People gravitate toward relationships that serve their A-R-C needs; social experiences that make people feel autonomous, related, and competent also promote stronger attachments.

Hypotheses

There is conceptual and empirical support for the view that a responsive object contributes to the creation of intense attachments, but much of this literature provides only indirect support, has been overlooked by marketing researchers, or has not been applied to the realm of human brands. Conceptually, Ryan and Deci (2000) argue that the satisfaction of A-R-C needs leads to the "development of secure attachments within specific relationships" (p. 326) because "proximal supports for basic psychological needs in any relationship ... play a crucial role in predicting feelings of attachment" (p. 262). Although Ryan and Deci focus on an interpersonal context, they propose that the A-R-C needs might be fulfilled in other ways, inviting the possibility that human brands qualify as responsive relationship partners.

Furthermore, La Guardia and colleagues (2000) asked participants to rate their levels of attachment security with respect to different interpersonal relationships. They also assessed the extent to which relationship partners satisfy participants' A-R-C needs. The results show that greater A-R-C feelings positively predict attachment security. However, because the study uses a composite measure of the A-R-C needs, it is not possible to determine the distinct contribution of each need in predicting attachment strength.

Autonomy. Empirical research suggests that there is a positive relationship between felt autonomy and attachment security in teenagers (Ryan and Lynch 1989). By providing an atmosphere in which autonomy is encouraged, parents facilitate the strengthening of attachment bonds. Prior work also documents a positive correlation between respondents' attachment and the extent to which their relationship partners make them feel autonomous in both interpersonal and human-brand contexts (Giles and Maltby 2004; Leak and Cooney 2001). Likewise, Schultz, Kleine, and Kernan (1989) argue that strong attachments are linked to the theme of individuation, reflecting a person's desire to feel autonomous and in control of the self. They report that many objects to which respondents are strongly attached show evidence of this theme. Taken jointly, these results suggest that the degree to which a relationship partner serves the need for autonomy should positively affect the intensity of a person's attachment to that partner.

H_1: The more a person perceives a human brand as fulfilling his or her autonomy needs, the more intensely attached he or she will be to the human brand.

Relatedness. Similar to Deci and Ryan's (2000) previously discussed arguments, Baumeister and Leary (1995) argue that the fulfillment of the need for relatedness by an object results in increased positive affect, such as feelings of attachment and love. In addition, prior literature has demonstrated empirically that many consumers who are strongly attached to particular objects show evidence of "integration," which, in part, refers to their preference to be connected and joined with others (Schultz, Kleine, and Kernan 1989, p. 360).

H_2: The more a person perceives a human brand as fulfilling his or her relatedness, the more intensely attached he or she will be to the human brand.

Competence. As discussed previously, it is not possible to distinguish the independent effects of each A-R-C need on attachment strength in all prior empirical work (e.g., La Guardia et al. 2000). This is important when considering the dearth of empirical work that addresses the independent contribution of competence. Conceptually, research suggests that attachments can develop for reasons tied to self-cultivation and self-development (Kleine and Baker 2004), themes that are at least consistent with the idea of a person feeling competent as a result of consuming a particular object. For example, elderly people in nursing homes surround themselves with special possessions that serve as reminders of a lost sense of proficiency (Rubinstein and Parmelee 1992). Similarly, consumers who partake in certain high-risk consumption activities can develop a sense of both mastery and emotional bonds in those contexts (Arnould and Price 1993; Celsi, Rose, and Leigh 1993), though it is unclear whether these bonds qualify as attachments because they may not be appropriately target specific. This research has not been tested in the context of human brands and provides only conceptual or indirect empirical support.

Evidently, the effect of the satiation of the need for competence on attachment is "a bit less straightforward" (La Guardia et al. 2000, p. 368). Although the satisfaction of both autonomy and relatedness needs is likely to occur primarily in emotionally significant relationships, a person's competence needs may be satiated in contexts that are unrelated to attachment, for example, at leisure, work, or elsewhere (Celsi, Rose, and Leigh 1993). Thus, if a husband fails to make his wife feel effective and capable (but does not actively make her feel ineffective or incapable), she may yet remain intensely attached to him because he satisfies her needs for relatedness and autonomy. Although theory implies that the fulfillment of the need for competence may qualify as responsiveness and thus produce intense attachments, prior empirical research is less conclusive. Thus, there is reason to predict that the linkage between competence and stronger attachments may be either positive or nonnegative.

H_{3a}: The more a person perceives a human brand as fulfilling his or her competence needs, the more intensely attached he or she will be to the human brand.

H_{3b}: The relationship between how a person perceives a human brand to be fulfilling his or her competence needs and intense attachments is nonnegative.

Study 1

Study 1 is designed to test H_1–H_3 directly. In addition, because prior literature has described how strong attachments are associated with satisfied, trusting, and committed relationships (Feeney and Noller 1996; Simpson 1990) and because these variables are also pivotal to marketing relationships (Fournier, Dobscha, and Mick 1998), Study 1 examines whether strong attachments positively predict the three variables. This endeavor extends prior work to a novel context, permits the testing of the independence of these constructs from attachment strength, and highlights the managerial implications of consumer brand attachments.

Method

One hundred sixty-four undergraduate volunteers, solicited in signs posted around a university campus, were paid $5 in exchange for filling out a survey that took most respondents approximately 15 minutes to complete. Young-adult respondents may be particularly well suited to studying the domain of human brands because they constitute a coveted target market for human-brand marketers and because it is common for them to develop attachments to human brands (Boon and Lomore 2001; Perse and Rubin 1989). Volunteers received one of two versions of an instrument whose preamble read as follows: "Many people think of themselves as being attached to certain celebrities or well-known figures (e.g., actors, athletes, singers ...). Think about a celebrity or well-known figure to whom you are very [not] attached." The two conditions (i.e., "very attached" and "not attached") were designed to create variance in responses on the dimension of attachment strength. Next, respondents answered a series of questions (in order) that mapped onto the dependent (i.e., attachment strength), independent (e.g., autonomy), and outcome (e.g., satisfaction) variables.

Measures

The measures for the A-R-C needs were taken from the work of La Guardia and colleagues (2000), who developed metrics appropriate for an interpersonal context (see the Appendix). Using two pretests, these measures were adapted to suit the current human-brand context, though minimal changes were actually required. Respondents were asked to indicate the extent of their agreement (anchored by 1 = "strongly disagree" and 7 = "strongly agree") with a series of statements about the human brand (denoted as "XYZ"). The items for relatedness ($\alpha = .81$) were as follows: "XYZ makes me feel cared about" and "I feel a lot of closeness with XYZ." Items for autonomy (α = .80) included "XYZ makes me feel controlled and pressured to be certain ways" (reversed) and "XYZ makes me feel free to be who I am." Items for competence (α = .70) included "Generally, XYZ makes me feel very capable and effective" and "XYZ makes me feel inadequate or incompetent" (reversed).

Previous research has argued that the degree to which people demonstrate separation distress is a good indicator of the strength of their attachment bonds (Berman and Sperling 1994). That is, with the real, imagined, or threatened separation from the object (e.g., death, loss), a strongly attached person will experience a negative emotional reaction that researchers can use as a measure of the intensity of the attachment bond. Prior research has advanced multi-item measures (Hazan and Shaver 1994; Hazan and Zeifman 1994) that were adapted and extended to the current context. The four items (anchored by 1 = "strongly disagree" and 7 = "strongly agree") show good reliability (α = .89): (1) "I feel better if I am not away from or without XYZ for long periods of time"; (2) "I miss XYZ when XYZ is not around"; (3) "If XYZ were permanently gone from my life, I'd be upset"; and (4) "Losing XYZ forever would be distressing to me."

Finally, the instrument included several metrics to gauge some of the potential positive qualities associated with people's stronger attachments and to provide further evidence that attachment strength is differentiable from prior constructs. Specifically, respondents were administered three metrics that were adapted from the perceived relationship quality scale (Fletcher, Simpson, and Thomas 2000), which taps the degree to which respondents assess their relationships to be satisfying, trusting, and committed. Although this scale has not been widely used in marketing, it shows excellent internal consistency and validity. The items (all anchored by 1 = "not at all" and 7 = "very much") for satisfaction (α = .97) were as follows: "How satisfied are you with the relationship?" "How content are you with the relationship?" and "How happy are you with the relationship?" Similarly, items for trust (α = .92) were as follows: "How much can you count on XYZ?" "How much do you trust XYZ?" and "How dependable is XYZ?" Finally, items for commitment (α = .98) were as follows: "How dedicated are you to the relationship?" "How committed are you to the relationship?" and "How devoted are you to the relationship?"

Results

There was an enormous range in the types of human brands that respondents listed, including actors, musicians, politicians, and media personalities. These figures represent a target of investment by, for example, movie studios, record labels, and political parties in terms of promoting the human brands and managing their brand associations. A manipulation check composed of an alternative measure of attachment strength (Thomson, MacInnis, and Park 2005) showed a significant difference between respondents in the weak- (M = 1.86) and those in the strong- (M = 4.05, p < .05) attachment conditions.

Results of a multivariate analysis of variance. The results of a multivariate analysis of variance provide initial support for the idea that A-R-C may be important in the production of stronger attachments. The mean of autonomy in the strong-attachment condition (M = 3.66, SE = .18) was higher than that in the weak-attachment condition (M = 1.98, SE = .17, p < .05). The mean of relatedness in the strong-attachment condition (M = 2.94, SE = .14) was significantly different from that in the weak-attachment condition (M = 1.30, SE = .13, p < .05), as was the mean of competence in the weak- (M = 2.71, SE = .18) versus the strong- (M = 3.67, SE = .19, p < .05) attachment condition.

Correlational analysis. A consideration of the correlation between the A-R-C needs and strong attachments marshals evidence of a significant linkage. Autonomy (r = .62, p < .05), relatedness (r = .70, p < .05), and competence (r = .44, p < .05) are each significantly correlated with attachment strength. In addition, the correlations among the A-R-C needs are positive (r's = .59–.75; see Table 1).

Full structural models. The remaining analysis is conducted with structural equation modeling (AMOS). It is customary to evaluate structural models in reference to alternative conceptualizations of the data and by comparing fit statistics. As such, the analyses included three models.

Model 1 is the hypothesized model, with the A-R-C indicators loading onto their respective factors, which are allowed to correlate and, in turn, predict attachment strength. Model 2 is the same as Model 1, except that the latent A-R-C needs are modeled as uncorrelated. Finally, Model 3 depicts all the indicators for the A-R-C needs loading onto a single latent factor.

In each model, the relationships between the indicators and the latent variables are positive and significant. Model 1

TABLE 1
Correlation Matrix of Study 1 Measures

	1	2	3	4	5	6	7
1. Autonomy	—	.75	.69	.62	.11	.58	.55
2. Relatedness		—	.59	.70	.30	.72	.69
3. Competence			—	.44	.12*	.49	.45
4. Attachment strength				—	.25	.66	.66
5. Satisfaction					—	.46	.33
6. Commitment						—	.68
7. Trust							—

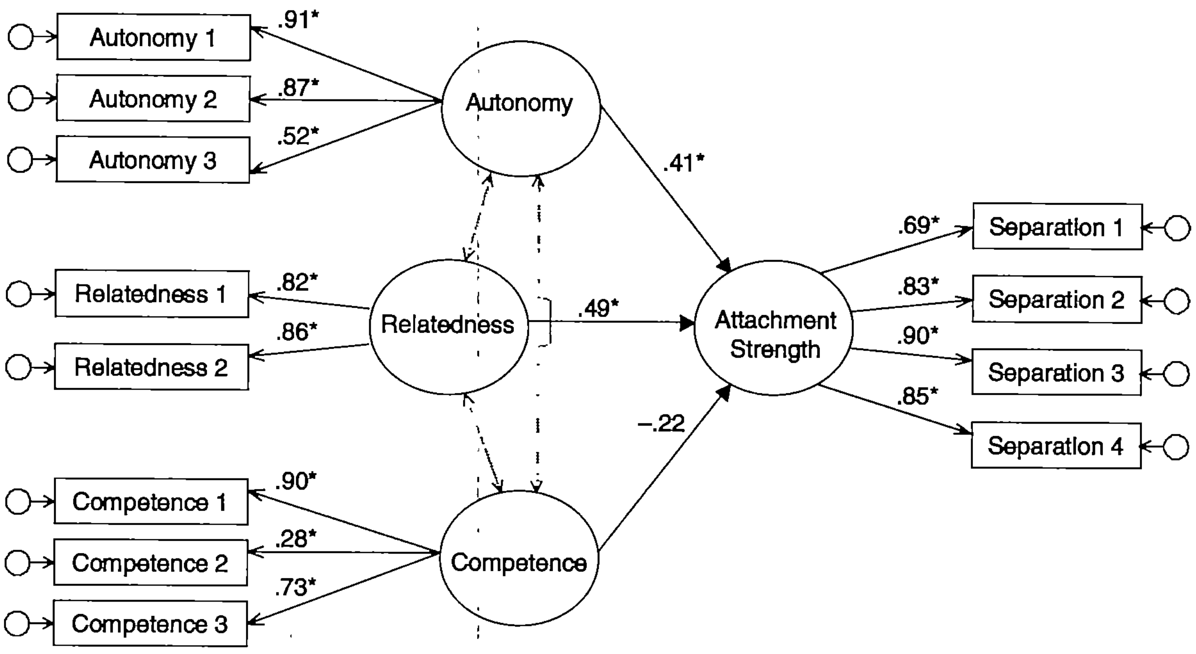
Notes: All *p*s < .05, except **p* = .14.

shows excellent minimum sample discrepancy divided by degrees of freedom (CMIN/DF = 1.92), comparative fit index (CFI = .96), and normed fit index (NFI = .92) statistics. Model 2, in which the A-R-C needs are uncorrelated, has universally unacceptable fit statistics and therefore is rejected. Specifically, the CMIN/DF statistic is inflated (6.04), and both the CFI (.77) and the NFI (.74) statistics are inferior. The fit statistics for Model 3 are mixed, with acceptable CMIN/DF (2.98) and CFI (.91) values but a marginal NFI value (.87). Because of its better fit and because it has a significantly different chi-square (χ^2 = 92.27, d.f. = 48) value from Model 3 ($p < .05$), Model 1 is superior (see Figure 1).

H_1 proposes that when a human brand fulfills autonomy needs, a person is more likely to become strongly attached to it. The path between autonomy and attachment strength

FIGURE 1
Summary of Study 1 Structural Models

A: Model 1



B: Model 2

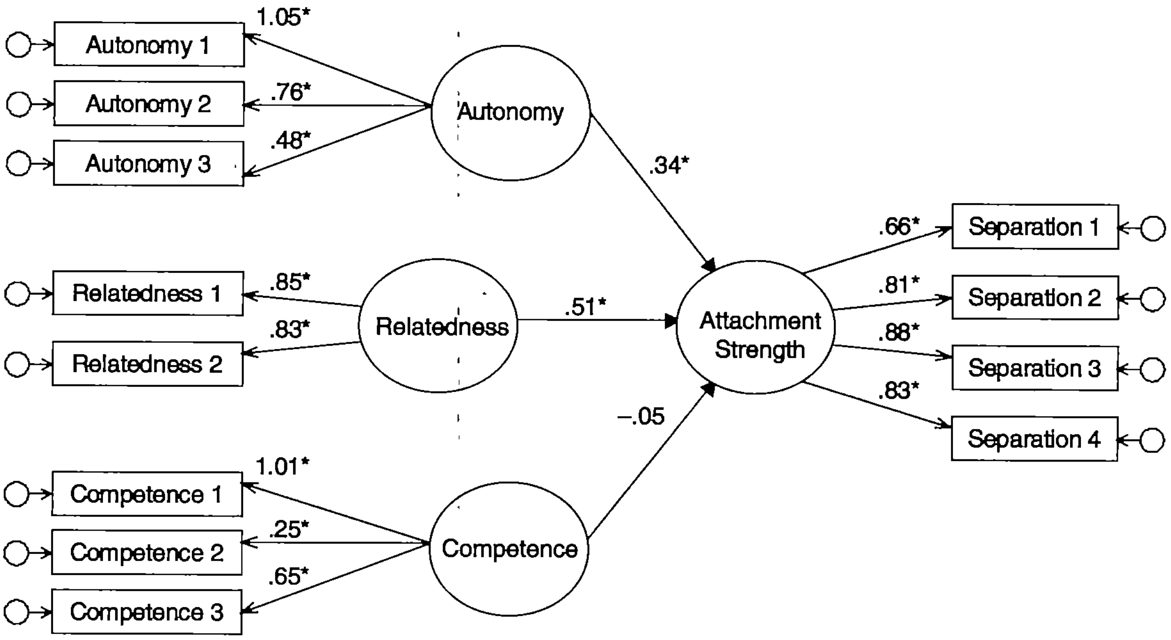
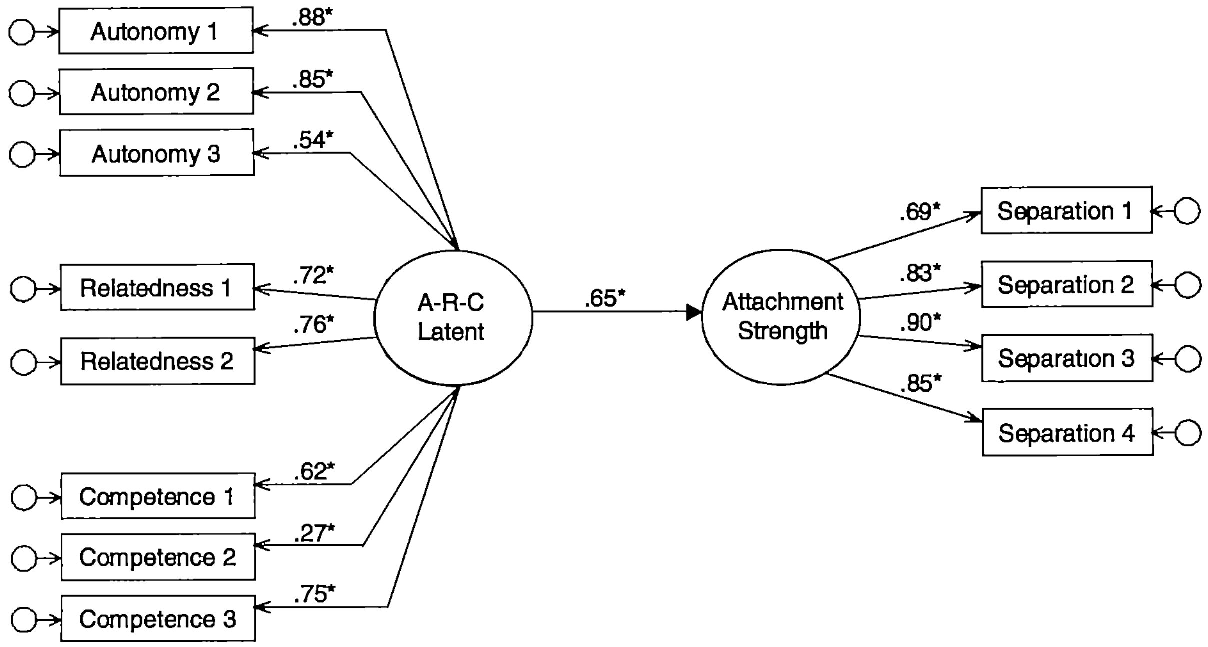


FIGURE 1
Continued

C: Model 3



*$p < .05$.

Notes: From Study 1, Model 1 shows all items for the independent variables loading onto their respective latent factors, which predict attachment strength (χ^2 = 92.26, d.f. = 48; CMIN/DF = 1.92; CFI = 96; and NFI = .92). Model 2 is identical to Model 1, except that the A-R-C latents are uncorrelated (χ^2 = 307.95, d.f. = 51; CMIN/DF = 6.04; CFI = .77; and NFI = .74). Model 3 shows all items for the independent variables loading onto one latent factor, which predicts attachment strength (χ^2 = 158 07, d.f. = 53; CMIN/DF = 2.98; CFI = .91; and NFI = 87). All path estimates are standardized and shown above the respective path.

in Model 1 is positive ($\gamma = .41$, $p < .05$), in support of H_1. The path between relatedness and attachment strength is also positive ($\gamma = .49$, $p < .05$), in support of H_2. Finally, H_{3a} is rejected, but H_{3b} is supported; there is an insignificant relationship between competence and attachment strength. This finding allows for the possibility that a human brand must not negatively affect competence for a strong attachment to exist.

Outcome variables. Although increasingly intense attachments may have positive qualities, these are not due to a lack of construct orthogonality. An exploratory factor analysis that included the attachment strength items and the satisfaction, trust, and commitment items was conducted (see Table 2). The results show a clean four-factor solution; no inappropriate loadings exceed .3, suggesting that these constructs are distinguishable.

Furthermore, stronger attachments are characterized by consumers' perceptions of their relationships with the human brand as satisfied, trusting, and committed. The correlation matrix (Table 1) shows that attachment strength is positively correlated with satisfaction (r = .25, $p < .05$), trust (r = .66, $p < .05$), and commitment (r = .66, $p < .05$). In addition, a multivariate analysis of variance of the differences in satisfaction, commitment, and trust between weak- and strong-attachment conditions suggests that there are benefits associated with strong attachments. In the strong- versus the weak-attachment conditions, respondents reported having greater satisfaction (M = 4.46, SE = .18 versus M = 3.14, SE = .19; F = 25.17, $p < .05$), commitment (M = 3.59, SE = .16 versus M = 1.66, SE = .17; F = 71.56, $p < .05$), and trust (M = 3.68, SE = .16 versus M = 1.90, SE = .17; F = 60.16, $p < .05$).

Discussion

The main contribution of Study 1 is that the fulfillment of autonomy and relatedness needs explains attachment strength, but the need for competence does not. Although theory and prior research suggests that an object that makes a person feel competent may bolster attachments (La Guardia et al. 2000; Rubinstein and Parmelee 1992), the results suggest that competence does not play a role in strengthening attachments to human brands. Although Table 1 reveals that competence is correlated with stronger attachments, its contribution is limited when the variance of autonomy and relatedness is taken into account. However, these same results suggest that the relationship between competence and attachment strength must be nonnegative. That is, an object that suppresses a person's competence needs may also diminish the strength of a person's attachment. This issue is considered again in the subsequent studies.

Although the results of Study 1 are encouraging, the choice of survey design dictates that some of the affective richness evident in consumers' attachments to human

TABLE 2
Study 1 Exploratory Factor Analysis with Relationship Constructs

	Factor			
Item	Satisfaction	Attachment Strength	Trust	Commitment
Satisfaction 1	.99			
Satisfaction 2	.95			
Satisfaction 3	.94			
Attachment Strength 1		.98		
Attachment Strength 2		.87		
Attachment Strength 3		.75		
Attachment Strength 4		.67		
Trust 1			.96	
Trust 2			.88	
Trust 3			.72	
Commitment 1				.97
Commitment 2				.93
Commitment 3				.91

Notes: Principal axis factoring with Promax rotation; all loadings >.3 are reported.

brands is lost (e.g., Rubin and McHugh 1987). It is clear from the literature that respondents use the same words to describe their attachments to human brands as they do to describe their strong interpersonal relationships (a result found by Fournier [1998] in reference to product brands). Study 1 builds primarily on the reliability and validity of the measures, whereas Study 2 employs a qualitative approach to corroborate the results of Study 1 and to flesh out some of the affective richness of consumers' attachments to human brands. In the process, Study 2 is intended to garner evidence that relationships characterized by strong attachments lie between and are comparable with both person-to-person (Hazan and Zeifman 1994) and product-brand (Fournier 1998) attachments.

Study 2

Study 2 is a qualitative study that evaluates evidence of the A-R-C needs in the context of human brands to complement the empirical results of Study 1. The use of qualitative procedures (Study 2) in conjunction with quantitative approaches (Studies 1 and 3) may be particularly well suited to research on attachments (Belk 1992) because, together, they richly describe the focal variable while statistically establishing the linkages among variables.

Method

Twenty-five undergraduate respondents were recruited on a university campus; each was paid $5 in exchange for answering questions about human brands in writing. No limitations were provided about which or what type of human brands to discuss; respondents were instructed to discuss any human brand ("celebrity or well-known person") to whom they felt "attached." In addition to documenting the emotional quality of these stronger attachments, the aim was to determine whether the responsiveness of human brands, revealed by the A-R-C needs, was evident in the context of these attachments because prior research is not conclusive on this issue. A series of questions were posed; for example, respondents were asked to tell the story of their first experience with the human brand, to describe their best memory involving the human brand, and to convey what the human brand meant to them. Thus, attachments were examined from the consumer's first-hand perspective on the basis of relatively unstructured descriptions of salient experiences with the brand (Fournier and Mick 1999). Two independent raters, who were blind to the purpose of the study, coded respondent feedback for evidence of strong attachments and the A-R-C needs. Disagreements between coders were resolved in discussions with the author.

Results

All responses showed evidence of the types of feelings typically associated with attachments of moderate to intense strength. In addition, 48% of responses indicate evidence of autonomy, and 60% show evidence of relatedness. Only one response shows evidence of the respondent experiencing feelings of competence. Incorporated into the subsequent sections are excerpts from the responses that tended to be extensive and thoughtful. In each, the respondent self-selected a human brand as the object of feelings of attachment. For example, one woman described a musical act in the following terms:

> I can't really remember when I started loving them. At least 15 years ago. I think probably I fell in love with the song "Rose Coloured Glasses," and then they played a rink in a town close to me and a friend and I went to see them. It was a great place to see them play because you could dance and get really close to the stage They had all these table and floor lamps on the stage. A really intimate setting, and there weren't that many people there. I think we paid $7 a ticket. They absolutely rocked. I've been addicted to them since then.

In this excerpt, the respondent describes a "loving" relationship that has endured for years. This pattern was common; many respondents referred to powerful emotions while describing facets of the relationships that were interpreted as evidence of specific A-R-C needs. For example, a respondent wrote the following about a band:

> I bought all of their albums without hearing them (I had only had the live one previous) and became obsessed with their talent. I went to see them for the first time ever and was so incredibly moved. I often found myself starring [sic] in awe at them during the show. I was even moved to tears on a few songs. After, I met Jim [lead singer] and got a picture taken with him, and I just couldn't believe at how nice he was. I had become from that experience, as other fans like to say, a Rodeohead. This was less than a year ago.

It is evident that this person's response to the band is affect laden. Furthermore, she now considers herself a "Rodeohead," a group of enthusiastic followers of the band (Blue Rodeo). Although this group appears to be fairly unstructured, many of its members gather in online forums (e.g., www.discussions.bluerodeo.com) to share stories, news, and opinions about the band. This excerpt demonstrates that a Rodeohead (and this respondent) enjoys social rewards that are linked to the band and that are interpreted as serving a need for relatedness. The following provides additional evidence of feelings of relatedness:

> Ah, now this is a better story. Another live show. This was my 3rd of 3 concerts;... it was in April of 2000 I think (I haven't gotten to a show since then—starving student and all) at the Grand Theatre.... It was really a fantastic night, my dad and I went together. He's a bit of a fan himself, though not quite as die-hard as myself. It's something that has sort of brought us together. We didn't have the greatest of seats because we bought our tickets late, but it didn't matter too much because the venue was fairly small. It was really a pretty magical evening. I found myself getting lost in the music now and then, and I just couldn't keep my eyes from glossing over.... Anyway, I've rarely felt so emotionally connected to anything.

Here, the respondent's relationship with the band is explained by the notion that it brings her closer to her father, a fairly clear representation of relatedness, even while her feelings of attachment are also directed at the band itself. A similar result with respect to feelings of relatedness is evident from the following excerpt in reference to a sports team, the Detroit Pistons:

> I used to attend games with my family as a child. I grew up watching Joe Dumars and Grant Hill all throughout high school. Basketball was a family event—a sanctuary—and I did my first internship with the company. The first game we went to was against the Portland Trailblazers—I was about 11 years old. Dennis Rodman grabbed 18 rebounds, and we cheered in the mid tier seats as they beat the Blazers by about 10 points. We had so much fun that game that going to the contests became a regular family event.

Other excerpts described people's adopting the mannerisms of various human brands, such as a male respondent who wrote, "I'm even a fan in their club. I own all of their CDs plus bootlegs. I even write and sign my name to reflect my love for them." This excerpt reveals evidence of relatedness ("fan in their club") and the type of self-endorsed behavior lying at the core of autonomy, reflected in the decision to model his signature after the style of print associated with the band's (Metallica) album covers.

Other passages showed similar modeling behavior that signifies a freedom to express the self. For example, a male respondent wrote about a band, "I strive to be like them because they haven't sold out and they are nice people who treat others well," and another respondent indicated that a particular performer (DMX) "inspires me to be a true man who will love my family and give back to my community. He isn't just a rapper. He's a legend and a great role model. He's one crazy hard-knocks guy, but he's also normal cuz [sic] he has kids and a wife and I respect that." Perhaps no example better captures the intersection of self-expression, self-endorsed behavior, and the feeling of freedom associated with autonomy as the following excerpt provided by a male respondent in relation to a musical band:

> As a musician myself (guitar, banjo, accordion, bass) I enjoy their folky-jazzy-rock vibe mixed with lyrics that hit home to me. Their songs are intelligent, well-crafted, and presented with impeccable taste. I think the fact that the band is accessible before and after many shows adds to their appeal. They sound like rock stars on stage—but act like normal guys off-stage. They provide songs that speak to me. They provide a middle-agreed [sic] family guy with the responsibilities of a family and mortgage and stressful career, an outlet to cut loose. They also inspired me to pick up my guitar and write songs again. My wife adores Jim Cuddy. Back in the mid-80's/early 90's the band was unique—now there is a whole industry (alt.country) dedicated to this blend of music. I am unique—as I follow a band that most everyone I meet has never heard of.

Finally, many respondents discussed objects they perceived as possessing a high degree of skill. For example, one respondent wrote about tennis player Martina Hingis having "a pretty solid game." Another respondent discussed Tori Amos as having "a spectacular singing voice, plays the piano masterfully, and is an amazing lyricist." One thought that Matthew McConaughey was "really good. We're talking Oscar here," whereas another described the World Wrestling Federation as follows: "The acting is superb and not just movie actors. These athlete/actors have to get each detail right because they are in front of a life [sic] audience 5–6 days a week." However, in all but one case, respondents attributed skill to the human brand. With one exception, none of the human brands appeared to make the respondents themselves feel competent. The single exception was a respondent who wrote the following about a band:

> I like to emulate what they have raised me on (so I feel). They were a positive role model for me in my teen years, providing a degree of sanity and a beacon of relief in a turbulent world I dealt with during that time. I admire all these traits in them and try to make them a part of my life in my own, different ways. Even when I meet a goal. I then set out for a new one to challenge me.

Whereas this excerpt shows evidence of self-expression ("make them part of my life"), the reference to being challenged is interpreted as fulfilling his need for competence. In contrast to the needs for autonomy and relatedness, there was only limited evidence for respondents' feeling attached to a human brand because of feelings of competence.

Discussion

The juxtaposition of feelings indicative of the A-R-C needs and feelings of attachment provides correlational support for the first and second hypotheses. The relative absence of feelings of competence experienced by the respondents does not contradict the results of Study 1. Both suggest that competence may be present but does not explain strong attachments in the context of human brands. However, there is also another possible explanation for this result, namely, that the nature of the human brands selected by respondents (primarily musical performers, athletes, and actors) or the age (18–25 years) of most respondents ruled out discussion of competence, which, in relation to other human brands (e.g., politicians, religious figures), might actually be linked to stronger attachments. However, it is clear that younger adults develop human-brand attachments (Boon and Lomore 2001; Cole and Leets 1999), suggesting that the choice of sample was appropriate (this issue is reexamined in Study 3). The instructions to respondents allowed them to discuss any human brand they desired. The predominance of actors, musicians, and sports figures in the sample corresponds to prior research that suggest that these are the most common categories of human brands that people typically describe (North, Bland, and Ellis 2005). Furthermore, respondents were clearly familiar with and actively contemplating the idea of skill and competence (e.g., Martina Hingis, Tori Amos), except that the attachment object was almost universally the target of such attributions. Study 3 further addresses the role of competence.

Study 2 presents evidence that respondents use the same words (e.g., love, adoration, emotional connection) to describe their attachments to human brands as they use in the context of other attachments (e.g., romantic). Thus, it would be premature to consider consumers' human-brand attachments trivial or unimportant; the meaningfulness and intense emotions ascribed to these relationships argue against this conclusion. Furthermore, although much of the literature cited in laying out the theoretical model originates in the interpersonal domain, the results suggest that the human-brand context is sufficiently similar to apply the A-R-C framework constructively.

In conjunction with Study 1, the results of Study 2 are encouraging, but additional support for the hypotheses and for the role of competence can be obtained by assessing attachment strength in a more natural environment. This is pursued in Study 3. In addition, consistent with the efforts of Study 2 to describe the phenomena more fully, Study 3 uses expanded measures of the A-R-C needs to ensure that the domain of these constructs is adequately captured.

Study 3

The hypotheses tested in Study 1 pertain to the role of the A-R-C needs in reference to human brands. Each is specified by self-determination theorists as potentially important to strengthening attachments. However, to address several alternative hypotheses, the analysis is expanded beyond the A-R-C needs. Specifically, there are other important human needs that are at the same high-order level of analysis as the A-R-C needs. For example, a recent publication that reviews and consolidates prior research on important human needs (see Sheldon et al. 2001) advances measures of various needs discussed in the overlapping research of scholars such as Maslow, Epstein, and Atkinson; these needs are referred to as (1) "self-actualization," (2) "pleasure–stimulation," (3) "money–luxury," (4) "self-esteem," and (5) "popularity–influence." This background is important in the current context because the extant literature seems to advance indirectly the more expansive view that servicing needs unrelated to the A-R-C needs may also bolster attachments. Although the articles cited subsequently support or are consistent with the view that the A-R-C needs are important, they also propose that through the course of an object's serving additional needs, a consumer may develop feelings that are emblematic of stronger attachments.

For example, Keller (2001) indicates that when a consumer gains feelings of fun, excitement, and self-respect from a brand, he or she may be more likely to become attached it. These feelings seem to mirror the needs for pleasure–stimulation and self-esteem (Sheldon et al. 2001). Therefore, a positive relationship between fulfillment of these two needs and increasingly strong attachments to human brands might be expected.

Similarly, Escalas and Bettman (2005) propose that a consumer will form self–brand connections when the brand helps him or her achieve goals, such as providing self-esteem, particularly when the brand is symbolic and when a person's in-group has already embraced the brand. This reasoning indicates that a person's needs for self-esteem, money–luxury (i.e., the brand's symbolic stature), and popularity–influence (i.e., the brand as a tool for a person to maintain position within a group) may be active in the process of attachment formation. In turn, this suggests that as an object services a person's needs for self-esteem, money–luxury, and popularity–influence, attachments become more intense.

Finally, Pimentel and Reynolds (2004) describe consumers who become more affectively committed to sports teams after undergoing a transcendent process of "sacralization," in which the team takes on meaningful status in the consumer's mind. This process seems consistent with Sheldon and colleagues' (2001) description of self-actualization, which refers to a person's identification of important aspects of life's meaning and purpose. Accordingly, it might be expected that the relationship between attachment strength and self-actualization, as well as the other basic needs, will be positive and significant.

> H_4: The more a person perceives a human brand as fulfilling his or her needs for (a) self-actualization, (b) pleasure–stimulation, (c) money–luxury, (d) self-esteem, and (e) popularity–influence, the more he or she will be intensely attached to that human brand.

This hypothesis does not undermine the role of the A-R-C needs in strengthening attachments. It is possible that these additional needs are complementary. Study 3 retests the hypotheses with a nonstudent sample in a more naturalistic setting and considers these five alternative needs in an effort to address the rival hypothesis that the satisfaction of *any* important need explains stronger attachments.

Method

Adult respondents were recruited by posting a letter of solicitation on Web sites organized as online communities for or by fans of various celebrities. These sites provide forums in which people gather electronically to share their thoughts and feelings about a celebrity with like-minded people. These sites were targeted because they are accessible and populated by people who are communicative and familiar with (though not necessarily attached to) at least one celebrity. Along other dimensions, such as age, geographic location, and gender, there is enormous heterogeneity in membership within and among sites.

A soliciting message was posted at each forum, which directed volunteers to send an e-mail to a university-affiliated address that was set up to for this study. To provide an incentive, the soliciting letter informed people that respondents who completed the survey would be entered into a random drawing for one of five prizes of $50 and would be eligible to receive a summary of the research. Those who sent an e-mail were directed to the location (i.e., a URL) of the survey, which was constructed for the Web using the nTreePoint software package. The study description, instructions, and order of presentation for the measures were nearly identical to those of Study 1, and respondents were randomly assigned to list a human brand to whom they either were or were not strongly attached.

Measures

The metric used to assess attachment strength was composed of the same four items used in Study 1 and showed good reliability (α = .93; see the Appendix). The A-R-C needs were each measured with the items employed in Study 1; additional items were adapted from the work of La Guardia and colleagues (2000) and Sheldon and colleagues (2001). Compared with Study 1, these showed improved internal consistency (αs = .83–.89).

The preamble for each metric that assessed needs read, "When I see, hear, think about or interact with this celebrity,..." The need for relatedness was measured with the two items from Study 1 plus four additional items: "I feel loved and cared about," "I often feel a lot of distance from this celebrity," "I feel like I am in contact with someone who cares for me and who I care for," and "I feel close and connected to a person who is important to me." The need for autonomy was measured with three items in addition to those used in Study 1: "I feel that my choices are based on my true interests and values," "I feel free to do things my own way," and "I feel that my choices express my 'true' self." The need for competence included three additional items: "I feel that I can successfully complete difficult tasks and projects," "I feel that I can take on and master hard challenges," and "I feel very capable in what I do."

Measures of other needs were included to assess the rival scenario that, in general, need fulfillment leads to strong attachments; these measures were adapted from the work of Sheldon and colleagues (2001). Respondents were asked in relation to each need the extent of their disagreement ("strongly disagree") or agreement ("strongly agree") with certain statements. Items for self-actualization (α = .95) were as follows: "I feel like I am 'becoming who I really am,'" "I feel a sense of deeper purpose in my life," and "I feel a deeper understanding of myself and my place in the universe." Pleasure–stimulation (α = .88) was measured with the following items: "I feel I am experiencing new sensations and activities," "I feel intense physical pleasure and enjoyment," and "I feel I have found new sources and types of stimulation for myself." The three items tapping money–luxury (α = .91) were as follows: "I feel I am able to buy most of the things I want," "I feel that I have nice things and possessions," and "I feel that I have plenty of money." The measure of self-esteem (α = .84) was composed of the following: "I feel I have many positive qualities," "I feel quite satisfied with who I am," and "I feel I have a strong sense of self-respect." Finally, people's need for popularity–influence (α = .85) was assessed by the following: "I feel that I am a person whose advice others seek out and follow," "I feel that I strongly influence other people's beliefs and behavior," and "I feel that I have a strong impact on what other people do."

Results

There were 107 respondents (26 males and 81 females), and they ranged in age from 18 to 60 years, with an average age of 33 years. Although gender and age were initially considered covariates in the subsequent analysis, neither predicted attachment strength and thus were dropped from the analysis. Although the instrument did not explicitly collect data on nationality, it was apparent from respondents' e-mail addresses and signatures that the sample was geographically diverse. For example, respondents appeared to be located in Australia, Canada, Hong Kong, New Zealand, South Africa, Spain, Sweden, Uruguay, and the United States.

There was an enormous range in the types of human brands selected by respondents, including actors, comedians, models, directors, radio personalities, writers, singers, athletes, musicians, politicians, and royalty. For example, respondents listed Ben Affleck, King Albert II, Lucinda Williams, Mario Lemieux, Viggo Mortensen, Michael Biehn, and Joan Crawford. Respondents also varied significantly in their popularity, activity level (e.g., retirement status), and profession.

Correlation. Each of the three A-R-C needs was positively correlated with the measure of attachment strength (see Table 3). The correlations between attachment and autonomy ($r = .37$, $p < .05$), relatedness ($r = .31$, $p < .05$), and competence ($r = .20$, $p < .05$) were all positive and significant. In general, the correlation between attachment strength and the other needs was weak or nonsignificant: self-actualization ($r = .16$, p = not significant [n.s.]), pleasure–stimulation ($r = .19$, $p < .05$), money–luxury ($r = .15$, p = n.s.), self-esteem ($r = .09$, p = n.s.), and popularity–influence ($r = .19$, $p < .05$).

Hierarchical regression. Using hierarchical regression allows variables entered in prior stages to account for variance in the dependent variable (i.e., attachment strength)

TABLE 3
Correlation Matrix of Study 3 Measures

Variable	1	2	3	4	5	6	7	8	9
1. Attachment strength	—	.38	.31	.20	.16*	.19	.15*	.09*	.19
2. Autonomy		—	.37	.60	.47	.20	.35	.33	.33
3. Relatedness			—	.52	.67	.52	.45	.42	.36
4. Competence				—	.56	.35	.46	.69	.45
5. Self-actualization					—	.76	.56	.48	.37
6. Pleasure–stimulation						—	.33	.38	.27
7. Money–luxury							—	.50	.44
8. Self-esteem								—	.66
9. Popularity–influence									—

Notes: All *p*s < .05, except *p = n.s.

before the independent contribution of variables entered in subsequent steps is assessed. Thus, it provides a relatively conservative test of the contributions of the variables entered last.

Considering the weak relationships between attachment strength and the non–A-R-C needs that were revealed by the previous correlation analysis, it might be expected that there is little or no relationship between an object's fulfilling these alternative needs and attachment strength. A consideration of the hierarchical regression with the alternative (i.e., non–A-R-C) needs entered in the first step and the A-R-C needs entered second shows that this is true.

In Regression 1 (see Table 4), in which the A-R-C needs are not entered, none of the alternative needs significantly predicts attachment strength. The coefficients for self-actualization ($\beta = .01$, p = n.s.), pleasure–stimulation ($\beta = .20$, p = n.s.), money–luxury ($\beta = .13$, p = n.s.), self-esteem ($\beta = -.06$, p = n.s.), and popularity–influence ($\beta = .10$, p = n.s.) are uniformly insignificant. In light of (1) the weak or insignificant correlations between attachment and the independent alternative needs and (2) the weak to moderate correlations among the alternative needs, it is unlikely that this result is due to multicollinearity.

Likewise, in Regression 2, which includes the A-R-C needs in the second step, none of the alternative needs is a significant predictor of attachment strength. Combined with the results from Regression 1, this demonstrates a lack of support for H_4. However, the links between two of the three A-R-C needs and the dependent measure are significant. Specifically, autonomy ($\beta = .31$, $p < .05$) and relatedness ($\beta = .29$, $p < .05$) are significant predictors, but competence ($\beta = -.03$, p = n.s.) is not. Thus, the results of Study 3 support H_1, H_2, and H_{3b}.

Discussion

Study 3 mirrors Study 1, insofar as both reveal that the satisfaction of autonomy and relatedness predicts strong attachments. In interacting with human brands, people who are made to feel capable of self-expression and self-governance and those who are made to feel a sense of closeness and intimacy are likely to develop strong attachments. In Studies 1 and 3 and in support of H_{3b}, the relationship between competence and attachment strength was not significant, suggesting that at least in a celebrity context, feelings of efficacy are not fundamental to promoting stronger attachments. Furthermore, competence and five additional needs were considered but were found to be insignificant in Study 3; this suggests that general need fulfillment does not explain the creation of strong attachments. Although other predictors of attachment strength may be identified in further research, the results suggest that autonomy and relatedness play an important role.

TABLE 4
Study 3 Hierarchical Regression Results

Step in Hierarchical Regression	Predictor	Regression 1 Standardized Beta	Regression 1 *p* Value	Regression 2 Standardized Beta	Regression 2 *p* Value
	(Constant)		.00		.01
Step 1	Self-actualization	.01	.94	–.31	.15
	Pleasure–stimulation	.20	.24	.26	.12
	Money–luxury	.13	.43	.14	.37
	Self-esteem	–.06	.72	–.13	.50
	Popularity–influence	.10	.52	10	.52
Step 2	Autonomy	—	—	.31	.03
	Relatedness	—	—	.29	.05
	Competence	—	—	–.03	.89

Notes: Dependent variable = attachment strength

General Discussion

Study 1 revealed that consumer attachments to human brands result in elevated levels of satisfaction, trust, and commitment. Study 2 revealed the affective richness of the experience for consumers attached to human brands. Study 3 offered additional support for the theoretical model using a separate sample. On the basis of this and other research (e.g., Ambler et al. 2002; Thomson MacInnis, and Park 2005), it is reasonable to conclude that attachment strength is separable from trust, satisfaction, commitment, and a variety of other constructs, such as favorable attitudes and brand loyalty, that have been linked to strong relationships. The promise of attachment strength to marketers is that the construct is predictive of these and other relationship features (e.g., Ahluwalia, Unnava, and Burnkrant 2001; Liljander and Strandvik 1995), is theoretically rigorous and well articulated, and is achievable through two means (i.e., creating feelings of autonomy and/or relatedness). Few, if any, relationship constructs appear to offer the same comprehensive benefits as attachment strength, which may provide a useful means of conceptualizing and an efficient metric for assessing the strength of consumer–brand relationships.

Managerial Implications

If consumers can be made to feel autonomous and related, organizations may be able to foster strong attachments and thus enjoy the benefits of superior relationships with consumers. In general, human brands that make consumers feel appreciated, empowered, and understood succeed in creating feelings of autonomy. Similarly, feelings of relatedness are likely to be produced when a human brand promotes acceptance, openness, and belonging.

General conditions. For an attachment to form, the consumer must initially be minimally attracted in some way to the human brand, whether socially or in some other manner (Boon and Lomore 2001). In other words, an attachment is unlikely to develop if the starting point of the relationship is characterized by intense negative feelings or thoughts. This suggests that managers must introduce a human brand to the world carefully and deliberately and choose a positioning that is appealing and sustainable over time.

Need for interaction. The importance of human brands interacting with consumers was underscored by the qualitative results from Study 2, in which one respondent spoke of having his picture taken with the human brand. Repeated interaction between a consumer and the human brand reduces uncertainty and provides the basis for an attachment to grow (Berman and Sperling 1994), but in the context of human brands, it is not just the quantity but also the quality of interaction that matters.

With direct interaction, consumers are more likely to view the human brand as accessible, increasing the opportunity for feelings of autonomy and relatedness to be promoted. Although "being seen" (e.g., public appearances) may be acceptable as a form of indirect interaction, more direct means offer a richer, more rewarding choice. Firms should focus on determining how to make their human brands approachable. For example, technological alternatives for direct interaction show potential not only because they expand the human brand's reach and exposure. If the human brand had a meaningful presence online through the use of blogs or chat rooms, attached consumers would gather to ask questions or to learn more about the human brand in his or her own words. An example of this direct approach is Mark Cuban, owner of the Dallas Mavericks basketball team, whose immense popularity is linked to his spending several hours daily answering e-mails from fans (Fielding 2005). The direct approach not only improves the quality of interaction but also may encourage feelings of autonomy and relatedness by reducing hierarchy, making consumers feel appreciated and understood, and reducing the prominence of economic motives by the human brand (Silvera and Austad 2004).

Interaction should also occur on a consistent basis (Beech, Chadwick, and Tapp 2000). Consumers may respond better to a human brand who has a routine that consumers may view as a guarantee of interaction. For example, Ellen DeGeneres is known for a daily segment on her show in which she invites the audience (both in the studio and watching on television) to dance with her. This is an especially astute move because of its regularity and because it simultaneously promotes feelings of autonomy (what DeGeneres calls "an expression of freedom") and closeness (dancing "with" a celebrity; see Elber 2005). Ensuring regular interaction between a consumer and a human brand helps create the conditions under which an attachment is likely to grow.

Role of authenticity. Human brands that are perceived as authentic are more readily embraced (Cole and Leets 1999; Perse and Rubin 1989). Research has elaborated on authenticity (Rose and Wood 2005), but what the term means in reference to a human brand remains an underresearched topic. However, several ideas are promising (see Tolson 2001). For example, authenticity is best developed slowly and deliberately, thus signaling a long-term view, and by avoiding perceived opportunism by the human brand, which can signal that the human brand has "sold out." Likewise, when human brands are seen publicly in a nonperforming role, they should not be "perceived as acting" (Tolson 2001, p. 445), because this may lead consumers to feel manipulated and to view the human brand as lacking credibility. In Study 2, one consumer spoke appreciatively of the band to which he was attached acting "like normal guys off-stage." Away from their trade, human brands should not be viewed as trying too hard to convey an image, or they will risk being viewed as "pretenders."

The role of competence. In all three studies, competence was not effective in bolstering attachments in the context of human brands. If the popular press is any indication, consumers can form attachments to human brands who lack any appreciable skills (e.g., Paris Hilton). However, human brands who make consumers feel incompetent are likely to weaken attachments or prevent them from forming in the first place. For example, when comedian Dennis Miller joined the Monday Night Football broadcast team, his addition was intended to halt a slide in ratings. However, on the show, Miller used indecipherable analogies, complex diction, and obscure references that alienated his audience and made them feel inadequate and unintelligent. In other words, many viewers were made to feel incompetent; this not only caused a weakening of any possible preexisting attachments but also prevented any new attachments from forming. Although consumers do not become attached to a human brand through feelings of competence, the results suggest that managers must nonetheless be careful not to permit the degradation of competence to be linked to their human brand.

However, caution must be taken in extending these results to other domains in which attachment and self-determination theories may provide a useful road map to building attachment bonds. For example, with respect to product functional brands that are intended to solve consumption-related problems (Park, Jaworski, and MacInnis 1986), consumers may experience feelings of competence as a result of brand use. Similarly, consumer relationships that implicate particular service providers, such as coaches or psychologists, may be explicitly designed to develop particular skills, and it is common for strong attachments to develop in such contexts (Colin 1996). However, only further empirical study of the issue can ascertain whether attachments in these contexts become strong because of the servicing of competence needs or whether competence is merely a correlate without any meaningful effect when the variance of autonomy and relatedness are taken into account.

Potential as endorsers. Marketing research that has considered human brands has tended primarily to address the choosing and measuring of the effectiveness of celebrity endorsers to enhance the appeal of a brand (Agrawal and Kamakura 1995). Although the current article does not directly address the use of human brands as endorsers, prior research and this article imply that human brands to whom consumers are attached may be unusually effective. One way to assess the appeal of a particular celebrity is the Q ratio, a commonly used but sometimes "misleading" measure of popularity indicated by the celebrity's familiarity and likeability (Knott and St. James 2002, p. 1428). Compared with the likeability dimension, an attachment is more significant and profound to an individual (Thomson, MacInnis, and Park 2005). Although consumers may *like* many celebrities, they will be strongly attached to only a few and be willing to sacrifice for and invest in these human brands to a greater extent.

In a related vein, research suggests that consumers who have positive feelings toward a particular brand are more likely to embrace a brand extension because of a more automatic and direct transfer of positive affect to the affiliated product (Yeung and Wyer 2005). This result occurs even when the extended brand has comparatively poor fit with the core brand, which implies that attachment may facilitate a highly effective transfer of feelings from the human brand to the sponsored object. In the context of celebrity endorsers, this suggests that there does not necessarily need to be a high degree of fit between an endorser and a particular brand if there is a strongly attached target market. A firm employing a human brand to whom consumers are strongly attached as an endorser may anticipate greater returns and insulation from transgressions (e.g., celebrity involvement in scandals; see Knott and St. James 2002) than a firm using a human brand who is merely well liked.

Limitations and Further Research

A result of these studies that has not yet been addressed is the overall weakness of attachments to human brands. In Study 2, respondents used language that suggested that their attachments were strong and emotionally significant. This result mirrors prior qualitative work on different types of consumer–brand relationships (Fournier 1998; Schultz, Kleine, and Kernan 1989). However, respondents in Study 1's strong-attachment condition reported attachment scores that, on average, hovered around the scale midpoint. The attachment scores in Study 3 were stronger than those in Study 1. Both studies used the same measure of attachment strength, but respondents in Study 3 may have been evangelical about the human brand they selected. After all, Study 3 respondents were recruited from online communities revolving around human brands. Although many of the participants in the low-attachment condition reported weak-attachment scores, (e.g., Orlando Bloom = 1.8/7, Matt Crane = 1.0/7), most respondents in the high-attachment condition reported scores greater than 6.5/7 (e.g., Ben Affleck, Julie Andrews, Bobby Valentine). These human brands seemed to play a more central role in the Study 3 respondents' lives than might otherwise be normally expected. Furthermore, although the same language might be used to describe interpersonal and consumption attachments (e.g., Study 2), relationships implicating the latter may be weaker in general. Human-brand attachments are probably best viewed as supplementary and subsidiary (Colin 1996). The implication for qualitative research on marketing relationships is that quantitative methods provide a useful perspective that may not be evident from qualitative approaches alone (Belk 1992).

Systematic research into what tactical choices are likely to promote autonomy and relatedness is needed. For example, do self-service models of service delivery or product customization promote autonomy by making consumers feel more self-expressive or self-governing? Do autonomy and relatedness need to be created from scratch, or can firms draw on existing feelings targeted elsewhere? Does possessing one or more strong consumption-related attachments mean that a person experiences elevated life satisfaction, or as alluded to by prior research (Kleine and Baker 2004), are there other, negative effects? Given the power of the strong attachments, answering these questions would be a worthwhile undertaking.

APPENDIX
Summary of Measures In Studies 1 and 3

	Study 1			Study 3		
Measure	**Items**	**α**	**M**	**Items**	**α**	**M**
Attachment strength	4	.89	2.33	4	.93	5.72
Autonomy	3	.80	2.69	6	.85	5.68
Relatedness	2	.81	1.90	6	.89	4.23
Competence	3	.70	3.10	6	.83	5.19
Self-actualization				3	.95	4.42
Pleasure–stimulation				3	.88	4.49
Money–luxury				3	.91	3.46
Self-esteem				3	.84	5.18
Popularity–influence				3	.85	4.45

Notes. All measures use seven-point Likert scales anchored by "strongly disagree" and "strongly agree."

REFERENCES

Agrawal, Jagdish and Wagner A. Kamakura (1995), "The Economic Worth of Celebrity Endorsers: An Event Study Analysis," *Journal of Marketing*, 59 (July), 56–63.

Ahluwalia, Rohini, H. Rao Unnava, and Robert E. Burnkrant (2001), "The Moderating Role of Commitment on the Spillover Effect of Marketing Communications," *Journal of Marketing Research*, 38 (November), 458–70.

Ambler, Tim, C.B. Bhattacharya, Julie Edell, Kevin Lane Keller, Katherine N. Lemon, and Vikas Mittal (2002), "Relating Brand and Customer Perspectives on Marketing Management," *Journal of Service Research*, 5 (1), 13–25.

Arnould, Eric J. and Linda L. Price (1993), "River Magic: Extraordinary Experience and the Extended Service Encounter," *Journal of Consumer Research*, 20 (June), 24–45.

Austin, James T. and Jeffrey B. Vancouver (1996), "Goal Constructs in Psychology: Structure, Process and Content," *Psychological Bulletin*, 120 (3), 338–75.

Baumeister, Roy F. and Mark R. Leary (1995) "The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation," *Psychological Bulletin*, 117 (3), 497–529.

Beech, John, Simon Chadwick, and Alan Tapp (2000), "Towards a Schema for Football Clubs Seeking an Effective Presence on the Internet," *European Journal for Sport Management*, (Special Issue), 30–50.

Belk, Russell W. (1992), "Attachment to Possessions," in *Place Attachment: Human Behavior and Environment*, Setha M. Low and Irwin Altman, eds. New York: Plenum Press, 37–62.

Berman, William H. and Michael B. Sperling (1994), "The Structure and Function of Adult Attachment," in *Attachment in Adults: Clinical and Developmental Perspectives*, Michael B. Sperling and William H. Berman, eds. New York: Guilford Press, 3–28.

Boon, Susan D. and Christine D. Lomore (2001), "Admirer-Celebrity Relationships Among Young Adults: Explaining Perceptions of Celebrity Influence on Identity," *Human Communication Research*, 27 (July), 432–65.

Carroll, Barbara A. and Aaron C. Ahuvia (2006), "Some Antecedents and Outcomes of Brand Love," *Marketing Letters*, forthcoming.

Celsi, Richard L., Randall L. Rose, and Thomas W. Leigh (1993) "An Exploration of High-Risk Leisure Consumption Through Skydiving," *Journal of Consumer Research*, 20 (June) 1–23.

Cole, Tim and Laura Leets (1999), "Attachment Styles and Intimate Television Viewing: Insecurely Forming Relationships in a Parasocial Way," *Journal of Social and Personal Relationships*, 16 (4), 495–511

Colin, Virginia L. (1996), *Human Attachment*. Philadelphia: Temple University Press.

Deci, Edward L. and Richard M. Ryan (2000), "The 'What' and 'Why' of Goal Pursuits: Human Needs and the Self-Determination of Behavior," *Psychological Inquiry*, 11 (4), 227–68.

Dietz, Park E., Daryl B. Matthews, Cindy Van Duyne, Daniel Allen Martell, Charles D.H. Parry, Tracy Stewart, Janet Warren, and J. Douglas Crowder (1991), "Threatening and Otherwise Inappropriate Letters to Hollywood Celebrities," *Journal of Forensic Sciences*, 36 (1), 185–209.

Edsall, Thomas B. and James V. Grimaldi (2004), "On Nov. 2, GOP Got More Bang for Its Billion, Analysis Shows," *Washington Post*, (December 30), A01.

Elber, Lynn (2005), "Ellen DeGeneres, Returning to Daytime TV, Still Revels in Freedom of Being Herself," *Sun Sentinel*, (September 6), (accessed December 12, 2005), [available at www.sun-sentinel.com].

Escalas, Jennifer Edson and James R. Bettman (2005), "Self-Construal, Reference Groups, and Brand Meaning," *Journal of Consumer Research*, 32 (3), 378–90.

Feeney, Judith A. and Patricia Noller (1996), *Adult Attachment*. London: Sage Publications.

Fielding, Michael (2005), "Mark, His Words: Marketing's Maverick on Blogs, Brands and Why Sales Is King," *Marketing News*, (November 1), 13–14.

Fletcher, Garth J.O., Jeffrey A. Simpson, and Geoff Thomas (2000), "The Measurement of Perceived Relationship Quality Components: A Confirmatory Factor Analytic Approach," *Personality and Social Psychology Bulletin*, 26 (3), 340–54.

Fournier, Susan (1998), "Consumers and Their Brands: Developing Relationship Theory in Consumer Research," *Journal of Consumer Research*, 24 (4), 343–73.

———, Susan Dobscha, and David Glen Mick (1998), "Preventing the Premature Death of Relationship Marketing," *Harvard Business Review*, 76 (January), 42–51.

——— and David Glen Mick (1999), "Rediscovering Satisfaction," *Journal of Marketing*, 63 (October), 5–23.

Giles, David C. and John Maltby (2004), "The Role of Media Figures in Adolescent Development: Relations Between Autonomy, Attachment and Interest in Celebrities," *Personality and Individual Differences*, 36 (4), 813–22.

Halpern, Jake (2005), "An Agent for Dead Celebrities," broadcast on *All Things Considered, National Public Radio*, (August 9), (accessed December 12, 2005), [available at www.npr.org].

Hazan, Cindy and Phillip R. Shaver (1994), "Attachment as an Organizational Framework for Research on Close Relationships," *Psychological Inquiry*, 5 (1), 1–22.

——— and Debra Zeifman (1994), "Sex and the Psychological Tether," in *Advances in Personal Relationships: Attachment Processes in Adulthood*, Vol. 5, Kim Bartholomew and Daniel Perlman, eds. London: Jessica Kingsley, 151–77.

Horovitz, Bruce (2003), "The Good, Bad and Ugly of America's Celeb Obsession," *USA Today*, (December 19), (accessed December 12, 2005), [available at www.usatoday.com].

Horton, Donald and Richard R. Wohl (1956), "Mass Communication and Para-Social Interaction: Observations on Intimacy at a Distance," *Psychiatry: Journal for the Study of Interpersonal Processes*, 19, 215–29.

Houran, James, Samir Navik, and Keeli Zerrusen (2005), "Boundary Functioning in Celebrity Worshippers," *Personality and Individual Differences*, 38 (1), 237–48.

Keller, Kevin Lane (2001), "Building Customer-Based Brand Equity: A Blueprint for Creating Strong Brands," *Marketing Management*, (July–August), 15–19.

Kleine, Susan Schultz and Stacey Menzel Baker (2004), "An Integrative Review of Material Possession Attachment," *Academy of Marketing Science Review*, 1, 1–39.

Knott, Cynthia L. and Melissa St. James (2002), "An Analytic Hierarchy Process Methodology Solution for Celebrity Endorser Decisions in Marketing," in *Decision Sciences Institute 2002 Annual Meeting Proceedings*. San Diego: Decision Sciences Institute, 1427–32.

La Guardia, Jennifer G., Richard M. Ryan, Charles E. Couchman, and Edward L. Deci (2000), "Within-Person Variation in Security of Attachment: A Self-Determination Theory Perspective on Attachment, Need-Fulfillment and Well-Being," *Journal of Personality and Social Psychology*, 79 (September), 367–84.

Leak, Gary K. and Ramie R. Cooney (2001), "Self-Determination, Attachment Styles and Well-Being in Adult Romantic Relationships," *Representative Research in Social Psychology*, 25, 55–62.

Lee, Michael (2005), "Iverson, 76ers Not Suited for Success," *Washington Post*, (November 2), E01.

Leets, Laura, Gavin de Becker, and Howard Giles (1995), "FANS: Exploring Expressed Motivations of Contacting Celebrities," *Journal of Language and Social Psychology*, 14 (March), 102–123.

Liljander, Veronica and Tore Strandvik (1995), "The Nature of Customer Relationships in Services" in *Advances in Services Marketing and Management*, Teresa A. Swartz, David E. Bowen, and Stephen W. Brown, eds. London: JAI Press, 141–67.

North, Adrian C., Victoria Bland, and Nicky Ellis (2005), "Distinguishing Heroes from Celebrities," *British Journal of Psychology*, 96 (1), 39–52.

Park, C. Whan, Bernard J. Jaworski, and Deborah J. MacInnis (1986), "Strategic Brand Concept/Image Management," *Journal of Marketing*, 50 (October), 135–45.

Perse, Elizabeth M. and Rebecca B. Rubin (1989), "Attribution in Social and Parasocial Relationships," *Communication Research*, 16 (February), 59–77.

Pimentel, Robert W. and Kristy E. Reynolds (2004), "A Model for Consumer Devotion: Affective Commitment with Proactive Sustaining Behaviors," *Academy of Marketing Science Review*, (accessed March 10, 2006), [available at http://www.amsreview.org/articles/pimentel05-2004.pdf].

Plunkett Research (2004), "Entertainment & Media Industry Overview," *Entertainment and Media Industry Statistics*, (accessed December 12, 2005), [available at www.plunkettresearch.com].

Rempel, John K., Michael Ross, and John G. Holmes (2001), "Trust and Communicated Attributions in Close Relationships," *Journal of Personality and Social Psychology*, 81 (July), 57–64.

Rindova, Violina P., Timothy Pollock, and Mathew Hayward (2006), "Celebrity Firms: The Social Construction of Market Popularity," *Academy of Management Review*, 31 (1), 279–92.

Rose, Randall L. and Stacy L. Wood (2005), "Paradox and the Consumption of Authenticity Through Reality Television," *Journal of Consumer Research*, 32 (1), 284–96.

Rubin, R.B. and M.P. McHugh (1987), "Development of Parasocial Interaction Relationships," *Journal of Broadcasting and Electronic Media*, 31 (3), 279–92.

Rubinstein, Robert L. and Patricia A. Parmelee (1992), "Attachment to Place and the Representation of the Life Course by the Elderly," in *Place Attachment*, Irwin Altman and Setha M. Low, eds. New York: Plenum Press, 139–63.

Ryan, Richard M. and Edward L Deci (2000), "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social-Development, and Well-Being," *American Psychologist*, 55 (1), 68–78.

——— and John H. Lynch (1989), "Emotional Autonomy Versus Detachment: Revisiting the Vicissitudes of Adolescence and Young Adulthood," *Child Development*, 60 (2), 340–56.

Schultz, Susan E., Robert E. Kleine, and Jerome B. Kernan (1989), "These Are a Few of My Favorite Things: Toward an Explication of Attachment as a Consumer Behavior Construct," in *Advances in Consumer Research*, Vol. 16, Thomas K. Scrull, ed. Provo, UT: Association for Consumer Research, 359–66.

Sheldon, Kennon M., Andrew J. Elliot, Youngmee Kim, and Tim Kasser (2001), "What Is Satisfying About Satisfying Events? Testing 10 Candidate Psychological Needs," *Journal of Personality and Social Psychology*, 80 (2), 325–39.

Sheth, Jagdish N. and Atul Parvatiyar (1995), "Relationship Marketing in Consumer Markets: Antecedents and Consequences," *Journal of the Academy of Marketing Science*, 23 (4), 255–71.

Silvera, David H. and Benedikte Austad (2004), "Factors Predicting the Effectiveness of Celebrity Endorsement Advertisements," *European Journal of Marketing*, 38 (11–12), 1509–1526.

Simon, Roger, Dan Gilgoff, and Terence Samuel (2004), "In the House of the Believers," *US News and World Report*, (August 9), (accessed March 10, 2006), [available at www.usnews.com].

Simpson, Jeffrey A. (1990), "Influence of Attachment Styles on Romantic Relationships," *Journal of Personality and Social Psychology*, 59 (5), 971–80.

Spake, Deborah F., Sharon E. Beatty, Beverly K. Brockman, and Tammy Neal Crutchfield (2004), "Consumer Comfort in Service Relationships," *Journal of Service Research*, 5 (4), 316–32.

Thomson, Matthew, Deborah J. MacInnis, and C. Whan Park (2005), "The Ties that Bind: Measuring the Strength of Consumers' Emotional Attachments to Brands," *Journal of Consumer Psychology*, 15 (1), 77–91.

Tolson, Andrew (2001), "'Being Yourself': The Pursuit of Authentic Celebrity," *Discourse Studies*, 3 (4), 443–57.

Towle, Angela Phipps (2003), "Celebrity Branding," *Hollywood Reporter*, (November 18), (accessed December 12, 2005), [available at www.hollywoodreporter.com].

Yeung, Catherine W.M. and Robert S. Wyer Jr. (2005), "Does Loving a Brand Mean Loving Its Products? The Role of Brand-Elicited Affect in Brand Extension Evaluations," *Journal of Marketing Research*, 42 (November), 495–506.

Raji Srinivasan

Dual Distribution and Intangible Firm Value: Franchising in Restaurant Chains

Dual distribution systems in which firms use vertical integration and market governance simultaneously are widely used across diverse marketing contexts (e.g., restaurants, retailing, industrial selling). A prominent example of dual distribution includes business format franchising, in which firms (i.e., the franchisors) license the operation of some of their units to franchisees while owning and operating some units themselves. Despite the widespread prevalence of dual distribution, there are few insights into its performance implications. In this article, the author examines the relationship between a firm's dual distribution strategy and its intangible value. Franchising in restaurant chains serves as the empirical context for the study. It is proposed that a firm's dual distribution strategy affects its intangible value both independently and jointly with a set of firm characteristics. The author considers the firm's age, scope of vertical integration, advertising, financial leverage, and financial liquidity as firm characteristics that influence the relationship between dual distribution strategy and intangible value. The author measures the firm's dual distribution strategy by the proportion of its franchised units to its total units and intangible value by its Tobin's q. The author estimates the proposed model using panel data on 55 publicly listed U.S. restaurant chains for the period 1992–2002. Unobserved firm heterogeneity is accommodated using latent class regression analysis. The results support a four-segment model. Dual distribution increases intangible value for some firms but decreases intangible value for others, both independently and in conjunction with other firm characteristics.

Dual distribution systems in which firms use vertical integration and market governance are widely used across diverse marketing contexts, including retailing, restaurants, and industrial selling. Although the nature of dual distribution systems varies by marketing context, in principle, dual distribution implies that the firm contracts with market agents to manage parts of its distribution system while owning and managing other parts of it. Given the widespread prevalence of dual distribution, some observers suggest that it has become the dominant channel design with important implications for firms' marketing strategies and performance (Anderson, Day, and Rangan 1997; Stern and Weitz 1997).

In this article, I examine the relationship between dual distribution strategy and intangible value using the context of franchising in restaurant chains. In what is known as business format franchising, the chains (i.e., the franchisors) own and operate some restaurant units and simultaneously license the operation of other restaurant units to market agents (i.e., the franchisees). Franchisors provide marketing and operations support to franchisees in return for royalty income, which is a fixed percentage of franchisees' revenues.

Dual distribution has been extensively studied by marketing scholars as hybrid channels, multiple channels, or plural governance. However, prior research has focused primarily on its various antecedents, including information asymmetry (Heide 1994, 2003), transaction costs (Anderson 1988; Anderson and Schmittlein 1984; Bergen, Dutta, and Walker 1992; Dahlstrom and Nygaard 1994, 1999; Rindfleisch and Heide 1997; Weiss and Anderson 1992), and uncertainty (Dutta et al. 1995; John and Weitz 1988). There are fewer insights into the performance implications of dual distribution. As Heide (2003, p. 27) notes, "the specific performance implications of plural systems remain unanswered;... establishing a link between particular governance approaches and outcome variables seems an important research priority."

Relating dual distribution strategy to intangible value is also valuable from a marketing metrics perspective (Bolton 2004; Rust et al. 2004). Recent developments in accounting suggest that a firm's accounting data alone cannot adequately explain its value because they do not capture its intangible assets (Amir and Lev 1996). Accordingly, there is interest in identifying value-relevant, nonaccounting measures, including business strategies that may affect shareholder value (e.g., Barth et al. 1998). Relating marketing strategies to firm value is especially pertinent because most marketing investments create intangible market-based

---

Raji Srinivasan is Assistant Professor of Marketing, Red McCombs School of Business, University of Texas at Austin (e-mail: raji.srinivasan@mccombs.utexas.edu). The author thanks Reuben Raj, Gary L. Lilien, Manoj Agarwal, Sundar Bharadwaj, Inge Geyskens, Natalie Mizik, Vithala Rao, and Rajendra Srivastava for providing useful input on previous versions of the article. The author also thanks the anonymous *JM* reviewers for their insightful and constructive comments.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

assets that are not reflected on balance sheets (Srivastava, Shervani, and Fahey 1998, 1999).[1]

Consistent with conceptual advances in the marketing metrics area (Srivastava, Shervani, and Fahey 1998, 1999), empirical research has identified a value-relevant role for various aspects of marketing strategy, including branding (e.g., Rao, Agarwal, and Dahloff 2004; Simon and Sullivan 1993), product management (Aaker and Jacobson 2001; Pauwels et al. 2004), and customer relationship management (e.g., Anderson, Fornell, and Mazvancheryl 2004). However, there are few insights into the relationship between distribution strategy, a key aspect of a firm's marketing strategy, and its value. An exception to this, however, is the work of Banerji and Simon (1991), who use survey data from 34 restaurants and report a positive relationship between the proportion of firms' franchised units to their company units and their intangible value. More recently, two studies (Geyskens, Gielens, and Dekimpe 2002; Lee and Grewal 2004) show that a firm's adoption of the Internet as a distribution channel positively influences its intangible value. These developments raise questions about the value relevance of distribution strategy.

A study relating dual distribution strategy to firm value is also important for marketing practice because of the increasing scrutiny of the financial accountability of marketing investments from various stakeholders (Rust et al. 2004). Testifying to the potential managerial significance of this issue, an empirical survey of restaurant chains, the industry context for this article, suggests an intriguing relationship between chains' dual distribution strategies and their intangible values (see Figure 1).

Notably, the scatter plot in Figure 1 indicates that restaurant chains using given dual distribution strategies have different intangible values. For example, in 2002, McDonald's Corporation and Panera Bread Company, which had similar dual distribution strategies (.72 of their system units were franchised), had intangible values of 1.290 and 5.300, respectively. Likewise, Benihana and California Pizza Kitchen, which had similar dual distribution strategies (.18 of their system units were franchised), had intangible values of .990 and 2.780, respectively. Similar patterns were observed in other industries in which dual distribution is prevalent (e.g., automotive repair services, hotels), suggesting a complex and perhaps contingent relationship between dual distribution strategy and intangible value. Insights clarifying this relationship will be useful to managers whose business strategies must be informed by their performance implications. In this article, I examine two research questions: Does a firm's dual distribution strategy affect its intangible value? and Do the effects of dual distribution on intangible value vary across firms?

Firms are urged to leverage synergy among elements of their marketing strategy to optimize resource allocation (Capon, Farley, and Hoenig 1990, pp. 1144–59; Gatignon and Hanssens 1987, p. 257). Surprisingly, however, prior studies relating marketing to firm performance have focused on one aspect of marketing, such as new product introductions (e.g., Bayus, Erickson, and Jacobson 2003) or brand management (e.g., Rao, Agarwal, and Dahloff 2004), without consideration of potential synergies among them. Indeed, the considerable variability in intangible values of firms for a given dual distribution strategy, as observed in Figure 1, suggests that a contingent relationship between dual distribution strategy and other firm characteristics affects intangible value.

Accordingly, I apply contingency theory (Hambrick 1983; Zeithaml, Varadarajan, and Zeithaml 1988), which proposes complementary effects of various aspects of a firm's profile, strategy, and resources on its performance, to relate a firm's dual distribution strategy to its intangible value. I propose that a firm's dual distribution strategy affects intangible value both independently and jointly with other firm characteristics. The firm's age, scope of vertical integration, advertising, financial leverage, and financial liquidity are considered characteristics that influence the effects of its dual distribution strategy on its intangible value.

I estimate the model using panel data on 55 publicly listed U.S. restaurant chains in the period 1992–2002, resulting in 394 firm years. Unobserved firm heterogeneity is accommodated using latent class regression analysis that derives latent segments (for a review, see Wedel and Kamakura 2000). I use the proportion of a chain's franchised units to the total number of its units in its distribution system as the indicator of its dual distribution strategy. I measure the chain's intangible value by its Tobin's q, a forward-looking, stock market–based measure (e.g., Bharadwaj, Bharadwaj, and Konsynski 1999).

The results support a four-segment model and indicate that a firm's dual distribution strategy influences its intangible value both independently and jointly with other firm characteristics. The effect of dual distribution on intangible value varies across firms in the four segments, in support of firm heterogeneity. The proposed model explains a firm's intangible value significantly better than its accounting data, in strong support of the value relevance of its dual distribution strategy. The model also predicts a firm's intangible value well, given its characteristics.

I organized the article as follows: In the next section, I define dual distribution and provide an overview of it. Then, I develop the hypotheses. In the subsequent sections, I describe the data and the model estimation procedure, and then I present the results. The article concludes with a discussion of the study's contributions, its limitations, and opportunities for further research.

Theory

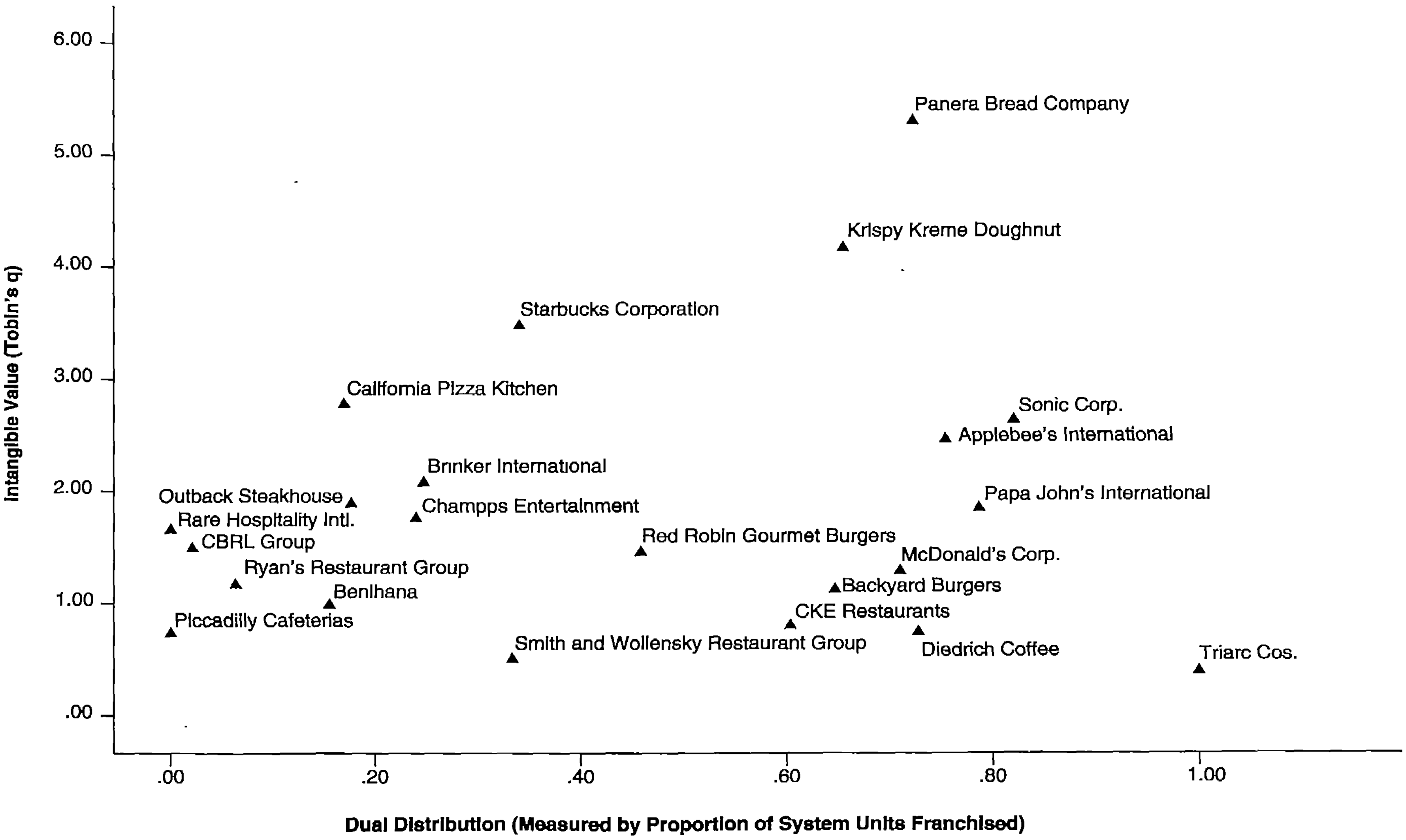
Definition of Dual Distribution

I conceptualize a firm's dual distribution as a hybrid system, involving the simultaneous use of vertical integration (known as "vertically integrated channels" or "company units" in franchising) and market governance (also called

---

[1] I use the terms "firm," "restaurant chain," and "chain" interchangeably to denote the restaurant chain and the terms "intangible value," "value," and "intangible firm value" to denote the firm's intangible value as measured by its Tobin's q.

FIGURE 1
Dual Distribution and Intangible Value In Restaurant Chains (2002)



"market-based channels" or "franchised units").[2] In the restaurant industry, most chains use dual distribution, though some (e.g., Planet Hollywood International, Lone Star Steakhouse & Saloon) use only vertically integrated channels and others (e.g., Noble Roman's, Triarc Companies) use only market-based channels.

I define the proportion of a chain's franchised units to the total number of its system units in its distribution system as the indicator of its dual distribution strategy for three reasons. First, this definition captures the chain's emphasis on market-based channels relative to its system size, reflecting the conceptualization of dual distribution as a hybrid system. Second, the definition also captures the chain's relative emphasis on the creation of intangible assets (i.e., franchised units, franchisee relationships, and knowledge) in its distribution system; this is consistent with the study's focus on firm value, which is achieved through intangible market-based assets (Srivastava, Shervani, and Fahey 1998). Third, the proportion of a chain's market-based channels to its system size is a strategic choice variable because chains first determine their overall distribution strategy (e.g., system size, location of units) and then determine which units should be franchised (Lafontaine 1992).

Overview of Dual Distribution

Vertically integrated channels and market-based channels embody different economic, structural, and managerial characteristics and serve different marketing functions (Dutta et al. 1995; Heide 1994, 2003). For example, firms own, operate, and manage vertically integrated channels through hierarchical, corporate control, whereas market agents own and operate market-based channels, which are managed primarily through provisions and incentives in a relational contract between the firm and its agents (Williamson 1975, 1985).

Marketing scholars have examined the conditions in which plural governance in general and dual distribution in particular lead to greater efficiency using the theoretical lenses of agency theory (Bergen, Dutta, and Walker 1992; Dutta et al. 1995) and transaction cost economics (Heide 1994, 2003; John and Weitz 1988; Rindfleisch and Heide 1997). The emergent consensus is that under some conditions, dual distribution eliminates problems associated with either exclusive vertical integration or market governance (for discussions of agency theory and transaction cost economics applicable to dual distribution, see Bergen, Dutta, and Walker 1992, pp. 13–16; Heide 1994, pp. 74–78; Rindfleisch and Heide 1997, pp. 32–43).

Advantages of Dual Distribution

Both vertical integration and market governance have distinctive economic and structural characteristics and marketing functions. Thus, the advantages and disadvantages of dual distribution arise from the aggregation of benefits and costs of each channel *and* the synergy across the dual channels.[3]

Greater control over operations in vertically integrated channels. In the presence of dual distribution, vertical integration enhances the credibility of the firm's contract termination safeguard, curtails agents' opportunism, and reduces the firm's vulnerability (Dutta et al. 1995). Vertical integration also provides a firm with greater control over its business processes, products, and profits (Heide 1994). Consistent with greater control over company units, the chief executive officer (CEO) of McDonald's stated, "We maintain a global base of company-operated restaurants to generate profits, link our interest with franchisees, develop management talent, gather research, and test ideas for better restaurant execution" (McDonald's Corporation 1995, p. 11).

Lower costs and higher returns in market-based channels. Market agents are an important source of scarce financial and managerial resources that are bundled efficiently in the contract (Alchian and Demsetz 1972; Grossman and Hart 1986; Jensen and Meckling 1976). As Anderson and Weitz (1986, p. 13) note, "administrative costs associated with external agents are quite low;… the 'invisible hand' of the market places substantial pressures on independent agents to provide a marketing activity efficiently."

For example, in restaurant chains, owner-managers of franchised units, with rights to residual profits from their units, are more motivated and work more efficiently than salary-based employees in company units, thus lowering operating costs. Chains monitor franchised units infrequently, relying instead on provisions in the contract, which lowers monitoring costs as well (Lafontaine and Shaw 2005). Furthermore, chains make no capital investment in franchised units, even as they earn income, which is set as a fixed percentage of franchisees' revenues. As the executive vice president at a leading fish-and-chips chain noted, "with franchise units, you receive a smaller margin but the return on investment approaches infinity because there is no investment" (Bradach 1998, p. 68).

Market-based channels are more responsive to market needs. Owner-managers of market-based channels follow complex strategies and are more responsive to market changes than are bureaucratic, vertically integrated channels (Yin and Zajac 2004). As a Pizza Hut franchisee noted (Bradach 1998, p. 127), "Franchisees have a faster reaction time. With all the layers in the company, it takes a long time to get things done." Because they are responsive to market needs, franchised units are an important source of innovations (Lewin-Solomons 1998). For example, three of

---

[2] The terms "multiple channels" and "dual channels" have also been used to denote the situation in which a firm owns and operates more than one channel (e.g., J.C. Penney has catalog operations, an online store, and bricks-and-mortar stores). I use the term "dual distribution" in this article to denote a situation in which one channel is owned and operated by the firm and the other channel is market-contracted to external agents. In addition, the term "hybrid" has been used in various ways in the literature. Some authors (e.g., Moriarty and Moran 1990) use it to describe a combination of institutional forms, as in this article, but others (e.g., Williamson 1991) use the term "hybrid" to describe a governance form that is intermediate between market governance and hierarchical governance.

[3] I gratefully acknowledge the guidance of two anonymous reviewers in the development of this section.

McDonald's most successful products—the Big Mac, Filet-O-Fish, and Egg McMuffin—were originally introduced by its franchisees (*BusinessWeek* 1997).

Market-based channels provide low-cost, low-risk access to new markets. Market-based channels provide firms with low-risk, low-cost access to markets that are not economically worthwhile to pursue with vertically integrated channels (Dahlstrom and Nygaard 1994). Indeed, franchised units are frequently located in remote, low-potential markets that, on average, achieve lower sales per unit than company units (Brickley and Dark 1987). Furthermore, the franchise contract is designed such that the upside benefit of increased revenues accrues to chains, whereas the downside risk of poor performance is borne solely by the franchisees. As Robert Nugent, the CEO of Jack in the Box, noted, "The economic advantage to franchising is that there is less risk to our income statement because we're not affected by cost pressures" (Bradach 1998, p. 68).

Synergy between dual channels. In addition to the distinct advantages of vertically integrated channels and market-based channels, synergy between the dual channels provides advantages (Lafontaine and Kaufmann 1994; Martin 1988). Chains leverage their experience in company units, over which they have control, to model responses in franchised units, over which they have much less control. Bradach (1997) calls this the "ratcheting process." In addition, Bradach (1997) finds evidence of two-way, "mutual-learning" processes between company units and franchised units. Testifying to synergy between the dual channels, the CEO of a restaurant chain noted (Bradach 1998, p. 2), "The chain [company units] gives you a system perspective, while franchisees give you a local perspective. We are constantly working to balance both perspectives. By having both company and franchised units, we are able to do that."

Disadvantages of Dual Distribution

Moral hazard problems in market-based channels. Because agents invest only in their business (i.e., the franchised unit), they experience higher risk than well-diversified investors and may invest suboptimally in their unit, creating moral hazard problems for the firm (Blair and Kaserman 1994; Fama and Jensen 1983). For example, franchisees may attract customers on the basis of the franchisor's reputation but deliver inferior products (e.g., Pizza Hut franchisees can reduce the quality and quantity of cheese in pizzas), thus profiting from such actions with limited negative consequences to them while hurting the firm's reputation. Such moral hazard problems are especially common when franchised units are located in remote markets characterized by high monitoring costs (Oxenfeldt and Kelly 1969).

Uncertainty about income from market-based channels. Restaurant chains that use dual distribution rely on market-based channels that are not under their control for a portion of their income, resulting in uncertainty about this income stream (Bradach 1997). Indeed, the forward-looking statements in chains' annual reports identify uncertainty in income from franchised units as an important risk factor (e.g., Papa John's International 2004). However, as I discussed previously, this risk is mitigated partly by vertical integration.

Hypotheses

The considerable variability in chains' intangible values for a given dual distribution strategy, as evident in Figure 1, suggests that a contingent relationship between dual distribution strategy and other firm characteristics affects intangible value. Accordingly, I use the contingency-based approach, which argues for a complementary effect of a firm's resources and strategy on its performance (Hambrick 1983; Zeithaml, Varadarajan, and Zeithaml 1988). Extending the contingency theory, I suggest that a firm's dual distribution strategy affects its intangible value both independently and jointly with its other characteristics. Thus, in addition to the main effect of dual distribution strategy on intangible value, I consider the effects of its interactions with the firm's age, scope of vertical integration, advertising, financial leverage, and financial liquidity on its intangible value. Note that though age, financial leverage, and financial liquidity are not marketing variables in the traditional sense, developments in agency theory (e.g., Bergen, Dutta, and Walker 1992) and transaction cost economics (e.g., Rindfleisch and Heide 1997) imply that they may influence the relationship between governance mode (i.e., dual distribution strategy) and firm performance.

Extending theoretical developments in market-based assets theory (Srivastava, Shervani, and Fahey 1998), I propose that under some conditions, defined by these interaction effects, the stock market will anticipate increasing cash flows, higher speed of cash flows, and lower risk of cash flows, which should increase the firm's intangible value. Under other conditions, the stock market will anticipate decreasing cash flows, lower speed of cash flows, and higher risk of cash flows, which should decrease the firm's intangible value. Because dual distribution offers both advantages and disadvantages, I hypothesize opposite effects when appropriate. The latent class regression approach, which I describe subsequently, accommodates opposite effects across firms in the different segments.

Main Effect of Dual Distribution

As I discussed previously, dual distribution offers several advantages, including lowering a firm's capital requirements, operating costs, monitoring costs, and risk. From a shareholder value perspective, market-based channels create relational, knowledge, and intellectual assets that increase future cash flow while reducing risk (Srivastava, Shervani, and Fahey 1998, 1999). From a real-options perspective (Bowman and Hurry 1993; Kulatilaka and Trigeorgis 1994), market-based channels are "options" investments that are useful for managing demand growth and uncertainty. Thus, an increasing proportion of market-based channels should increase a firm's intangible value.

However, agency theory accounts suggest some disadvantages of dual distribution as well, including potential underinvestment and free riding by market agents (Bergen, Dutta, and Walker 1992). Thus, an increasing proportion of market-based channels in such firms may decrease their future cash flows and increase the uncertainty of these cash flows, thus negatively affecting the firm's intangible value.

Given these opposing arguments, I expect that there is a positive effect of the proportion of market-based channels on intangible value for some firms and a negative effect for others. Thus:

> H_{1a} (H_{1b}): The greater the proportion of market-based channels in a firm that uses a dual distribution strategy, the higher (lower) is its intangible value.

Dual Distribution and Firm Age

Over time, most high-potential markets are covered by a firm's existing distribution system so that its expansion may be limited to smaller, remote markets, which are more efficiently served by market-based channels (Lafontaine and Shaw 2005). In such a situation, the stock market may view an increasing proportion of market-based channels as a cost-effective way to grow, thus strengthening the firm's cash flows and increasing its intangible value.

However, other arguments from the strategy and franchise literature suggest an opposite effect. The life cycles of firms influence their resource profiles and strategies (Penrose 1959). Franchising is important in the early years of a firm when its resources are scarce (Carney and Gedajlovic 1991). Over time, as resources increase, franchising may become less important, resulting in the well-documented reversion to company ownership of franchised units in older firms (e.g., Lafontaine and Kaufmann 1994; Oxenfeldt and Kelly 1969). Thus, the older the firm and the greater the proportion of its market-based channels, the lower is its future cash flows, the higher is the risk of these cash flows, and the lower is its intangible value. Given these opposite effects,

> H_{2a} (H_{2b}): The interaction between the proportion of market-based channels in a firm that uses a dual distribution strategy and its age increases (decreases) its intangible value.

Dual Distribution and Scope of Vertical Integration

In the current context, the scope of a firm's vertical integration is the proportion of the firm's sales from its vertically integrated channels to its total sales (John and Weitz 1988). Several factors suggest a positive effect of the interaction between the scope of a firm's vertical integration and dual distribution on its intangible value.

A firm that depends more on its vertically integrated channels for sales will have higher fixed investments in its integrated channels, offering a credible termination safeguard against agents' opportunism and improving both the agents' and the system's performance (Heide 1994). Furthermore, a firm that depends more on its vertically integrated channels for sales will have higher marketing expenditures (Lafontaine 1992; Lafontaine and Shaw 2005), creating positive marketing spillovers for its market-based channels and, again, strengthening its performance.

In addition, when a firm has an increasing proportion of sales from its vertically integrated channels and a large proportion of market-based channels, its market-based channels will have lower sales per unit than its vertically integrated channels, suggesting that they are covering smaller markets with lower potential. In such a situation, the firm's market-based channels are a cost-effective way to reach these smaller markets without attendant capital investment and risk. Thus, the interaction between the scope of a firm's vertical integration and an increasing proportion of market-based channels may strengthen its future cash flows, reduce the risk of these cash flows, and increase intangible value.

> H_3: The interaction between the proportion of market-based channels in a firm that uses a dual distribution strategy and the scope of the firm's vertical integration increases its intangible value.

Dual Distribution and Advertising

Empirical studies indicate that the effects of advertising on performance are industry specific, varying in magnitude, timing, and persistence (Comanor and Wilson 1967; Ravenscraft and Scherer 1982). The interest here is in the effect of the interaction between a firm's advertising and the proportion of its market-based channels on intangible value.

The firm's and the market agents' incentives are closely aligned when brand equity, which is consistent with high advertising, is high (Arrow 1985; Mathewson and Winter 1984). For example, when a firm has high advertising, its market agents may be motivated to perform better to extract higher rents from the higher brand reputation, improving system performance. Furthermore, when a firm with high advertising also has a high proportion of market-based channels, its advertising investment may be efficiently recovered over the larger market, increasing the firm's cash flows and reducing their uncertainty. Thus, the higher the firm's advertising and the higher the proportion of its market-based channels, the higher is its intangible value.

However, other arguments suggest the opposite. Agency theorists (Brickley and Dark 1987) argue that market agents of firms with higher advertising may have an incentive to free ride on the firm's reputation, especially in remote markets with high monitoring costs. For example, franchisees of well-known brands located in remote markets may reduce product quality, inferring that such actions bear no negative consequences for them or for the firm. However, such actions can negatively affect the firm's product perceptions, reputation, and performance. Thus, when a firm with high advertising also has a high proportion of market-based channels, the increased moral hazard threat may lower its future cash flows and increase the risk of these cash flows, lowering the firm's intangible value. Given these opposite effects,

> H_{4a} (H_{4b}): The interaction between the proportion of market-based channels in a firm that uses a dual distribution strategy and its advertising increases (decreases) its intangible value.

Dual Distribution and Financial Leverage

In the current context, financial leverage is the extent to which a firm uses debt to finance its assets (Jensen and Meckling 1976). A firm's financial leverage affects its strategy and performance through agency problems between the firm and its bondholders and stockholders (Bradley, Jarrell, and Kim 1984). As I discuss next, these agency problems may result in an effect of the interaction between a firm's financial leverage and dual distribution strategy on its intangible value.

Stock and bond markets cannot differentiate between "good-quality" and "bad-quality" firms. Thus, good-quality firms with high financial leverage signal their quality with actions that are too costly for bad-quality firms (Myers and Majluf 1984). For example, good-quality firms may invest in fixed assets, increasing the security for their bondholders and reducing their cost of capital (Myers 1977).

By the reverse logic, when a firm with high financial leverage has a high proportion of market-based channels, an inexpensive source of financial resources (Carney and Gedajlovic 1991), the bond market may interpret this as a signal of low funds, inferring the firm's inability to service its debt, lower its debt rating, and increase its cost of capital (Lang, Ofek, and Stulz 1996). In such a situation, the stock market may expect the firm's future cash flows to decrease, even as its risk increases, thus negatively affecting its intangible value. Thus:

> H_5: The interaction between the proportion of market-based channels in a firm that uses a dual distribution strategy and its financial leverage decreases its intangible value.

Dual Distribution and Financial Liquidity

Firms vary in their financial liquidity, or acid ratio (i.e., the extent to which current assets cover current liabilities); a higher acid ratio suggests a larger proportion of current assets than current liabilities. Financial liquidity affects a firm's strategy and performance through investment constraints (e.g., Bond and Meghir 1994), creating agency problems between the firm and its shareholders (Fazzari, Hubbard, and Petersen 1988; Hsiao and Tahmiscioglu 1997). A high acid ratio suggests cash availability to fund investments with superior earnings potential. Indeed, the stock market prefers a lower acid ratio to a higher one because it views current assets as less productive than fixed assets (Myers and Rajan 1998).

When a firm has a high proportion of market-based channels and a high acid ratio, the stock market may infer that the firm is forgoing income generation opportunities by not investing in vertically integrated channels. Thus, the stock market may expect a firm's increasing liquidity and an increasing proportion of its market-based channels to lower its intangible value. Thus:

> H_6: The interaction between the proportion of market-based channels in a firm that uses a dual distribution strategy and its financial liquidity decreases its intangible value.

Method

Data

I used unbalanced panel data of publicly listed U.S. restaurant chains for the period 1992–2002, which I obtained from various sources, including Technomic (an industry market research firm), Compustat, and the firms' annual reports. The data included 55 firms observed for 394 firm years. Some firms entered after 1992 or exited before 2002, contributing fewer firm years. The average number of firm years was 7.164 (minimum = 1 year; maximum = 11 years).

The firms in the study constituted 28% of publicly listed restaurant chains. The profile of firms in the study did not differ from the population of firms in Compustat (N = 198) in terms of profit ($t_{(\text{profitability})} = 1.012$, not significant [n.s.]), but they were larger ($t_{(\text{sales})} = 1.912$, $p < .10$; $t_{(\text{employees})} = 1.987$, $p < .05$). The firms in the study are representative of the population, though they are slightly larger. This is not surprising, because larger firms are favored in public data sources.

Measures

Tobin's q is the measure of a firm's intangible value. Chung and Pruitt's (1994) method was used to calculate Tobin's q:

$$q = \frac{\text{MVE} + \text{PS} + \text{DEBT}}{\text{TA}}, \quad (1)$$

where MVE is the closing price of shares at the end of the financial year × number of common shares outstanding; PS is the liquidation value of outstanding preferred stock; DEBT is the sum of book value of inventories, long-term debt, and current liabilities less current assets; and TA is the book value of total assets.

As I discussed previously, the proportion of the restaurant chain's franchised units to the total number of its system units, a continuous measure bounded between 0 (only vertically integrated channels) and 1 (only market-based channels), was the indicator of its dual distribution strategy. The trends in dual distribution for the period 1992–2002 suggested variability in dual distribution across chains and over time. The firm's age was measured by the number of years since its incorporation. The scope of vertical integration was measured by the proportion of the firm's sales from its company units to its total sales. The firm's sales from its company units and its total sales were obtained from Technomic and Compustat, respectively. The firm's advertising was measured by its advertising stock, using advertising expenditures reported in Compustat.[4] To account for the persistence of advertising, advertising stock was measured using one-, two-, and three-year lags with annual discount rates of .45 (Hirschey and Weygandt 1985). Through the use of information criteria, the three-year lag was identified as the appropriate lag.[5] The firm's financial leverage was measured as the ratio of its long-term debt to its total assets (Jensen and Meckling 1976), and the firm's financial liquidity was measured as the ratio of its cash, marketable securities, notes receivable, and accounts receivables to its current liabilities (Schmidgall, Hayes, and Ninemeier 2002). Table 1 contains the descriptive statistics and the correlation matrix of the measures.

---

[4] None of the firms in the study reported research and development expenditure.

[5] I thank an anonymous reviewer for this suggestion. The Schwarz Bayesian information criterion (BIC) for the four-segment model (reported subsequently) without profiling variables is 784.682, and it is 786.029 for models with advertising stock with a one-year and two-year lagged term for the advertising stock, respectively, compared with 778.354 for the model with a three-year lag for the advertising stock. Thus, the three-year lagged term for advertising stock was selected as the appropriate lag term. In addition, following the suggestions of a reviewer, I reestimated the model with advertising stock adjusted for the firm's sales. In general, the results were consistent, but the model with the advertising adjusted for sales performed worse (BIC =

The correlations were within acceptable limits (highest correlation = –.438 between firm age and financial leverage). Following the work of Belsley, Kuh, and Welsch (1980), I assessed potential threats from multicollinearity. The variance inflation factor and condition numbers were lower than 10 (average = 1.238; maximum = 3.456) and 15 (average = 1.875; maximum = 2.893), respectively, suggesting that multicollinearity is not a threat to the validity of the study's findings.

Latent Class Regression

To account for unobserved firm heterogeneity, I used latent class regression that simultaneously classifies observations into latent segments and estimates regression models within each segment (see Wedel and Kamakura 2000). The latent class regression approach directly identifies latent segments on the basis of the inferred relationship between a response variable (i.e., Tobin's q) and the set of explanatory variables (i.e., dual distribution and related interactions) within each homogeneous group (i.e., a given latent segment). An outline of latent class regression models appears in the Appendix.

Results

Correction for Sample Selection Bias

Some firms (n = 20) were not observed for the entire period. If the data were not missing randomly and the model was estimated with the observed data, the parameter estimates may be biased (Heckman 1979). I corrected for potential sample selection bias by including Lee's lambda (Inverse Mill's Ratio), which I obtained from a Heckman selection model in the regression model for Tobin's q.[6]

I first estimated the Probit selection model, including the firm's age, sales, and return on assets as factors that influence its exit from the data set. The results support the selection model ($\chi^2 = 20.342$, d.f. = 1, $p < .01$). The firm's age (b = –.063, $p < .01$), size (b = –.005, $p < .01$), and return on assets (b = –.965, $p < .01$) had negative effects in the selection model. Consistent with the liabilities-of-newness arguments in organizational sociology (Brüderl and Schussler 1990), newer, smaller, and less profitable firms were more likely to exit.

Model Estimation and Selection

Following the procedure outlined in the Appendix, I estimated two sets of latent class regression models (1) with the restaurant chain's class (i.e., whether it was casual dining, quick service, and limited service) and the number of brands in its portfolio (e.g., Yum! Brands had four brands, including KFC, A&W, Pizza Hut, and Taco Bell) as profiling variables and (2) with only a constant term (i.e., no restaurant characteristics) as the profiling variable. I included all main effects of explanatory variables used to create the interaction effects to ensure complete model specification. I also included the firm's return on its shareholders' equity, measured as the ratio of net income to shareholders' equity, and a quadratic term of the proportion of its market-based channels as control variables in the regression model.[7]

To determine the number of latent segments, I used the Bayesian information criterion (BIC) to compare the model with s segments with a model with s + 1 segments $\forall$ s = 1, 2, ..., until the model fit stopped improving (Schwarz 1978). The four-segment model, without the restaurant characteristics term for profiling, outperformed models with

TABLE 1
Descriptive Statistics and Correlations (N = 394)

Variable	M (SD)	1	2	3	4	5	6	7
1. Tobin's q	1.548 (1.046)	1.000						
2. Dual distribution (proportion of market-based channels)	.390 (.326)	.025	1.000					
3. Firm age (years)	26.980 (18.000)	–.052	.229**	1.000				
4. Scope of vertical integration	.654 (.562)	–.105*	–.079	–.116*	1.000			
5. Advertising stock (in millions of dollars)	58.132 (152.693)	.120*	.273**	.347**	–.222**	1.000		
6. Financial leverage	.259 (.237)	–.185**	.123*	–.438**	–.199**	.207**	1.000	
7. Financial liquidity	.741 (.978)	.185**	.221**	–.231**	.013	–.099*	–.225**	1.000

*p < .05
**p < .01.

801.932) than the proposed model (BIC = 778.354). I obtained similar results using the consistent Akaike's information criterion.

[6] For details on the statistics of sample selection correction, see Heckman (1979).

[7] I included the quadratic term for dual distribution, following the suggestion of an anonymous reviewer.

one, two, three, and five segments both with and without profiling variables, as well as the four-segment model with profiling variables.[8] The entropy of separation for the selected four-segment model was .972, which compares well with the results in the literature (Wedel and Kamakura 2000).

Tests of Hypotheses

Table 2 contains the results of the four-segment model with the constant term for profiling. First, Lee's lambda significantly affected a firm's intangible value in two segments (Segment 2: b = –.379, $p < .01$; Segment 4: b = 1.405, $p < .01$), in support of the correction for sample selection bias. There were several unhypothesized main effects of the explanatory variables. Specifically, the negative effect of advertising stock on intangible value was counterintuitive (Segment 2: b = –.001, $p < .01$; Segment 4: b = –.031, $p < .01$). I examine this result in detail subsequently.

As I predicted in H_{1a} and H_{1b}, respectively, the results support both the positive (Segment 2: b = .694, $p < .01$) and the negative (Segment 1: b = –.321, $p < .05$) effects of the firm's proportion of market-based channels on its intangible value. In addition, the square of the proportion of market-based channels had an unhypothesized positive effect on intangible value in Segment 1 (b = 1.190, $p < .01$). As I predicted in H_{2a} and H_{2b}, respectively, the results support the positive (Segment 1: b = .022, $p < .01$) and negative (Segment 2: b = –.034, $p < .01$; Segment 4: b = –.091, $p < .01$) effects of interaction between the firm's age and the proportion of its market-based channels on its intangible value. As I predicted in H_3, the interaction between the scope of a firm's vertical integration and the proportion of its market-based channels had a positive (Segment 2: b = .634, $p < .05$) effect on its intangible value. Somewhat unexpectedly, the interaction between the scope of a firm's vertical integration and the proportion of its market-based channels also had a negative effect on its intangible value (Segment 1: b = –.413, $p < .01$). I examine this result in detail subsequently.

The results also support H_{4a} and H_{4b}, which predict positive (Segment 1: b = .005, $p < .05$; Segment 2: b = .005, $p < .01$) and negative (Segment 4: b = –.073, $p < .01$) effects of the interaction between a firm's advertising stock and the proportion of its market-based channels on its intangible value. As I predicted in H_5, there was a negative effect of the interaction between a firm's financial leverage and the proportion of its market-based channels on its intangible value (Segment 2: b = –1.634, $p < .01$; Segment 3: b = –7.207, $p < .05$). Finally, as I predicted in H_6, there was a negative effect of the interaction between financial liquidity

TABLE 2
Results for the Model Relating a Firm's Dual Distribution to Its Intangible Value (N = 394)

Variable	Segment 1[a]	Segment 2	Segment 3	Segment 4
Intercept	.853 (.049)***	1.393 (.071)***	1.891 (.254)***	.593 (.592)
H_{1a}/H_{1b}: Dual distribution[b]	–.321 (.135)**	.694 (.149)***	.564 (.605)	–1.110 (1.009)
H_{2a}/H_{2b}: Dual distribution × firm age	.022 (.007)***	–.034 (.009)***	.028 (.028)	–.091 (.031)***
H_3: Dual distribution × scope of vertical integration	–.413 (.160)***	.634 (.253)**	.100 (.781)	– 140 (.206)
H_{4a}/H_{4b}: Dual distribution × advertising stock	.005 (.002)**	.005 (.001)***	–.003 (.010)	–.073 (.025)***
H_5: Dual distribution × financial leverage	–.275 (.471)	–1.634 (.449)***	–7.207 (3.215)**	7.175 (5.218)
H_6: Dual distribution × financial liquidity	–.507 (.072)***	–.138 (.153)	–.367 (.474)	–2.144 (1.043)**
Firm age	–.004 (.002)**	.006 (.003)**	.009 (.004)**	–.038 (.014)***
Scope of vertical integration	.001 (.000)***	–.001 (.000)*	–.001 (.000)***	.001 (.000)***
Advertising stock	.001 (.002)	–.001 (.000)***	–.002 (.003)	–.031 (.007)***
Financial leverage	.947 (.135)***	–.661 (.142)***	–3.150 (.864)***	–3.019 (1.384)**
Financial liquidity	.323 (.033)***	–.192 (.050)***	.199 (.179)	.585 (.405)
Dual distribution²	1.190 (.358)***	–.695 (.498)	–3.151 (2.657)	1.328 (1.486)
Return on shareholders' equity	.035 (.025)	.013 (.013)	.884 (.348)**	–.582 (.235)**
Lee's lambda	.011 (.049)	–.379 (.113)***	.111 (.111)	1.405 (541)***
Segment size (number of firms)	36% (20)	29% (16)	20% (11)	15% (8)

*$p < .10$.
**$p < .05$
***$p < .01$.
[a]Parameter estimates and standard errors.
[b]Dual distribution is measured by the proportion of market-based channels in the firm.

---

[8] For the model formulation with restaurant class and number of brands as profiling variables, the BIC is 1131.655 for the one-segment model, 861.773 for the two-segment model, 846.560 for the three-segment model, 822.564 for the four-segment model, and 808.100 for five-segment model. Similarly, for model formulation with only the constant term as the profiling variable, the BIC is 1131.655 for the one-segment model, 843.380 for the two-segment model, 828.113 for the three-segment model, 778.354 for the four-segment model, and 786.839 for the five-segment model. I obtained similar results using the consistent Akaike's information criterion.

and the proportion of a firm's market-based channels on its intangible value (Segment 1: $b = -.507$, $p < .01$; Segment 4: $b = -2.144$, $p < .05$). Thus, the results strongly support the hypotheses.

Latent Segment Profiles

In this section, I report the post hoc segment profiles of the four segments. The results of an analysis of variance (see Table 3) suggest that there are significant differences across firms in the four segments. The following is a summary of the Scheffé multiple comparisons of means of variables across the four segments (Pedhazur 1997):[9]

- *Intangible value.* Firms in Segment 4 had higher intangible value than firms in the other segments ($p < .01$), firms in Segment 3 had higher intangible value than firms in Segments 1 and 2 ($p < .01$), and firms in Segment 2 had higher intangible value than firms in Segment 1 ($p < .05$).
- *Return on assets.* Firms in Segment 1 had a much lower return on assets than other firms ($p < .01$).
- *Size.* Firms in Segment 2 were larger in terms of system size, total assets, and sales than other firms ($p < .01$), whereas firms in Segment 4 were smaller in system size than other firms ($p < .01$).
- *Number of brands.* Firms in Segments 3 had more brands than other firms ($p < .01$).
- *Advertising stock.* Firms in Segment 2 had higher advertising stock than other firms ($p < .01$), whereas firms in Segment 4 had lower advertising stock than other firms ($p < .05$).
- *Financial leverage.* Firms in Segment 4 had much lower financial leverage than firms in the other segments ($p < .05$).

I integrate the segment profiles with the regression results to explain the findings across the segments. Firms in Segment 1 (e.g., Backyard Burgers, Nathan's Famous, Carrols Corporation) performed poorly in this period. For these firms, the proportion of market-based channels had a U-shaped relationship ($b_{(\text{linear})} = -.321$, $p < .05$; $b_{(\text{quadratic})} = 1.190$, $p < .01$) on their intangible values. Perhaps the stock market infers that given their poor overall performance, these firms have limited ability to manage market-based channels, though notably, this negative effect disappears at very high proportions of market-based channels. For firms in Segment 1, the scope of vertical integration has a positive main effect on intangible values ($b = .001$, $p < .01$), but this effect is reversed in conjunction with the proportion of their market-based channels ($b = -.413$, $p < .01$). The positive main effect of vertical integration suggests that for these poorly performing firms, increasing their sales from the company-based channels reassures the stock market of their future cash flows. Conversely, the negative main effect of the proportion of market-based channels, both independently and in conjunction with the scope of vertical integration, implies that the stock market may view these poorly performing firms as not being well positioned to manage market-based channels. However, as these firms aged, the proportion of their market-based channels positively affected their intangible values ($b = .022$, $p < .01$), suggesting that, over time, market-based channels are an effective mechanism for these firms to achieve growth. Likewise, as firms' advertising stock increased, the proportion of market-

TABLE 3
Summary of Results Across the Four Latent Segments

Variable	**Segment 1**	**Segment 2**	**Segment 3**	**Segment 4**
Intangible value (22.192**)[a]	1.048	1.300	1.903	2.735
Return on assets (8.345**)	.006	.018	.016	.028
System size (units) (5.698**)	1671	2229	1064	643
Total assets (in millions of dollars) (10.082**)	398.451	1450.611	443.613	343.658
Sales (in millions of dollars) (9.178**)	653.684	1091.030	630.852	518.781
Number of brands (3.184*)	1.400	1.375	2.545	1.375
Advertising stock (in millions of dollars) (9.231**)	42.363	83.832	32.557	23.388
Financial leverage (6.792**)	.323	.315	.205	.155
Exemplar Members	**Backyard Burgers, Nathan's Famous, Carrols Corp.**	**IHOP Corp., McDonald's Corp., Wendy's International**	**Brinker International, CBRL Group, AFC Enterprises**	**California Pizza Kitchen, Panera Bread Company, Diedrich Coffee**
Dual distribution (H_{1a}/H_{1b})(+/–)	–/+(quadratic)	+	n.s.	n.s.
Dual distribution × firm age (H_{2a}/H_{2b}) (+/–)	+	–	n.s.	–
Dual distribution × scope of vertical integration (H_3) (+)	–	+	n.s.	n.s.
Dual distribution × advertising (H_{4a}/H_{4b}) (+/–)	+	+	n.s.	–
Dual distribution × financial leverage (H_5) (–)	n.s.	–	–	n.s
Dual distribution × financial liquidity(H_6) (–)	–	n.s.	n.s.	–

*$p < .05$.
**$p < .01$.
[a]Figures in parentheses are the overall test of means across the four segments ($F_{(3, 51)}$).

---

[9] There are no significant differences across firms in the four segments on other variables.

based channels positively affected their intangible values ($b = .005$, $p < .05$), suggesting that market-based channels also enable effective leverage of advertising stocks.

Firms in Segment 2 (e.g., IHOP Corporation, McDonald's Corporation, and Wendy's International) were large firms with large advertising stocks. For these firms, the proportion of market-based channels positively affected their intangible values ($b = .694$, $p < .01$), suggesting that perhaps because of market saturation, market-based channels that tap small and low-potential markets are an efficient way to grow. The effects of the interactions between proportion of market-based channels and both advertising ($b = .005$, $p < .01$) and scope of vertical integration ($b = .634$, $p < .05$) positively affected firms' intangible values. Notably, the main effects of both advertising ($b = -.001$, $p < .01$) and scope of vertical integration ($b = -.001$, $p < .10$) for these firms are negative. For large firms with large advertising stocks, perhaps the ceiling effects of advertising result in the negative main effect. However, the stock market may view these large firms as being able to leverage their large advertising stocks effectively over a larger market using market-based channels, resulting in the observed positive effect of the interaction between advertising stock and the proportion of market-based channels on intangible values. The negative main effect of the scope of vertical integration and the positive effect of the interaction between the proportion of market-based channels and the scope of vertical integration suggest that the stock market views expansion through market-based channels as the most effective way for these large firms to achieve growth.

Firms in Segment 3 (e.g., Brinker International, CBRL Group, AFC Enterprises) were small firms with a larger portfolio of brands than other firms. Notably, most hypothesized effects of dual distribution on intangible value were not supported for these firms (the one exception was the interaction between financial leverage and proportion of market-based channels: $b = -7.207$, $p < .05$). Overall, the intangible values of these firms appeared to be impervious to their dual distribution strategies, perhaps because the stock market views their brand portfolio diversification as a sufficient hedge against market uncertainty.

Firms in Segment 4 (e.g., California Pizza Kitchen, Panera Bread Company, Diedrich Coffee) were very small firms with low advertising stocks and low financial leverage. For these firms, there was no main effect of the proportion of market-based channels ($b = -1.110$, n.s.) on their intangible values, but advertising had a negative main effect ($b = -.031$, $p < .01$) and the scope of vertical integration had a positive main effect ($b = .001$, $p < .01$) on their intangible values. However, there were negative effects of the interactions between the proportion of market-based channels and advertising stock ($b = -.073$, $p < .01$), age ($b = -.091$, $p < .01$), and financial liquidity ($b = -2.144$, $p < .05$) on their intangible values. The negative effect of advertising stock on intangible value, both independently and in conjunction with dual distribution, is intriguing. Indeed, after the large firms in Segment 2, the firms in Segment 4 have the highest proportion of advertising stock to system size. I conjecture that the stock market may perceive these firms as having adequate advertising stock and thus need to achieve system growth. In addition, for these firms, dual distribution has a negative effect on intangible value. Integrating this pattern of findings suggests that perhaps the stock market views these firms as good custodians of intangible value and rewards them for organic growth through company-based channels (i.e., lower emphasis on dual distribution).

Additional Analyses

Model comparisons. First, I compared the proposed model with a model with *only* the accounting data of financial leverage, liquidity, and return on shareholders' equity. The hypothesized model (BIC = 778.354) outperformed this baseline model (BIC = 823.358). Second, I compared the hypothesized model with a model that included the firm's age, scope of vertical integration, advertising stock, and financial accounting data, but I *excluded* dual distribution and the related interaction effects. Again, the hypothesized model (BIC = 778.354) outperformed the model that excluded dual distribution (BIC = 798.567). Third, I compared the proposed model with a model that included main effects of dual distribution and other explanatory variables, but I *excluded* the interaction effects. Again, the hypothesized model (BIC = 778.354) outperformed the model that excluded the interaction effects (BIC = 810.187). Thus, the proposed model with the inclusion of dual distribution and related interactions explained firm value better than models that included only accounting data or models that excluded dual distribution; this shows strong support for the value relevance of the dual distribution strategy.

Predictive validity. I evaluated the model's predictive ability using a jackknifing technique, holding out a target firm, reestimating the model on other firms, and then using the estimated parameters to predict the target firm's Tobin q. I computed the mean absolute deviation (MAD), defined as $(1/NT)\Sigma_{t_i=1}^{T_i}\Sigma_{i=1}^{N}|\text{Tob\_q}_{oit} - \text{Tob\_q}_{pit}|$, where Tob_q_{oit} and Tob_q_{pit} denote observed and predicted Tobin's q for firm i in year t, N denotes the number of firms, and T_i denotes the number of years for each firm i, respectively. The MAD using the proposed model was .473, compared with a MAD of .713 with the average value, representing an improvement of 34%. Thus, the proposed model predicts the firm's intangible value well.

Discussion

Despite the growing importance of dual distribution channels in practice, there are few insights into their performance implications. In this article, I explore the relationship between dual distribution strategy and intangible value. I conclude with a discussion of the theoretical contributions, managerial implications, and limitations and opportunities for further research.

Theoretical Contributions

Marketing channels. By relating dual distribution strategy to intangible value, this article addresses the call for research on the performance implications of plural governance in general and dual distribution in particular (Heide 2003). The findings indicate that a firm's dual distribution strategy significantly affects its intangible value. Furthermore, the differential main and interaction effects of dual distribution across the four latent segments suggest firm heterogeneity in the value relevance of dual distribution.

For some firms (Segments 1 and 2), an emphasis on dual distribution contributes both independently and jointly with other firm characteristics to their intangible values. For others (Segment 3), their intangible values are mostly impervious to their dual distribution. Still others (Segment 4) are punished by the stock market for emphasizing dual distribution, albeit in conjunction with other firm characteristics. This evidence of firm heterogeneity in the value relevance of dual distribution empirically validates prior theoretical developments (e.g., Bergen, Dutta, and Walker 1992; Heide 1994) that indicate both advantages and disadvantages of dual distribution. The generalizability of the contingent nature of the performance effects of dual distribution to other performance metrics (e.g., profit, market share) emerges as an area for further research.

Importantly, from a shareholder value perspective, dual distribution does not offer unilateral benefits to all firms, as implied by some observers (e.g., Bradach 1998; Moriarty and Moran 1990). Instead, it is contingent on their other characteristics, including marketing and financial strategies. This finding, which stresses synergies between aspects of firms' distribution and business strategies, represents a useful extension to marketing strategy research in general and to marketing channels research in particular, which, for the most part, has not examined such synergies.

Marketing metrics. This article's findings also contribute to the marketing metrics literature, in which there is a gap with respect to the relationship between a firm's distribution strategy (dual distribution, in particular) and its performance. By providing strong evidence of the value relevance of dual distribution strategy, this study takes a first step toward addressing this gap.

In a departure from prior studies in both marketing and accounting literature, which have used a functional silo-based approach (i.e., using either marketing or accounting measures, respectively) to study the value relevance of strategy, this article's findings support interactions between aspects of marketing and finance strategies. These interactions suggest a previously unexamined potential interchangeability between marketing and financial strategies in maximizing intangible value. For example, a firm with high financial liquidity (firms in Segment 1 and Segment 4) can decrease its emphasis on dual distribution to increase its shareholder value. Exploring the boundaries of such interchangeability across marketing and other aspects of business strategies (e.g., finance, accounting, human resources) is a worthwhile area for further research, given the growing importance of the financial accountability of marketing (Rust et al. 2004).

Furthermore, the value relevance of dual distribution suggests that intangible, market-based channels not accounted for on firms' balance sheets but rather on their market agents' balance sheets significantly affect their intangible values. This finding implies the existence of boundaries beyond firms' balance sheets that are critical for an accurate assessment of their intangible values. The demarcation of such extended boundaries for inventorying value-relevant, intangible, market-based assets emerges as an important area for further research.

Finally, the unhypothesized negative effect of advertising stock on intangible value (Segment 2 and Segment 4) is counterintuitive and merits discussion. For the large firms (Segment 2), perhaps ceiling effects on their large advertising stocks produce this negative effect. For the small firms (Segment 4), perhaps the stock market views advertising as an exploratory, risky, value creation strategy and an emphasis on market-based channels as an exploitative, value appropriation strategy (March 1991). These findings, combined with those of Mizik and Jacobson (2003), who report a value creation role for research and development and a value appropriation role for advertising, imply that whether the stock market views a given strategy as creating or appropriating value may be context specific.

Managerial Implications

The study's findings also generate useful implications for marketing practice. For example, the significant and different main and interaction effects of dual distribution on firms' intangible value across the four latent segments indicate that the value relevance of dual distribution is contingent on other firm characteristics. This clarifies the puzzling relationship between chains' dual distribution strategies and their intangible values reported in Figure 1. For example, in 2002, the differences in intangible values between McDonald's Corporation (ticker: MCD: Tobin's q = 1.290) and Panera Bread Company (PNRA: Tobin's q = 5.300), both with similar dual distribution strategies (.72 of their system units are franchised), may be partly due to the differences in their characteristics, including age (MCD = 47 years; PNRA = 15), system units (MCD = 31,108; PNRA = 478), advertising stock (MCD = 1088.112; PNRA = 13.445 [in million of dollars]), financial leverage (MCD = .400; PNRA = .000), and financial liquidity (MCD = .489; PNRA = 1.348), which influence the relationship between their dual distribution and their intangible values.

Given the increasing pressures on marketing executives to demonstrate the financial accountability of marketing and the limited prior research on the financial accountability of distribution strategy, the value relevance of dual distribution is both a novel and a useful insight. Marketing managers can use the current findings to convince various stakeholders (e.g., top management and finance executives, some of whom may be skeptical of the financial accountability of marketing) of the value relevance of marketing in general and dual distribution strategy in particular. Managers can use the study's specific findings to develop investor communications programs to increase their firms' intangible value. In addition, given the model's predictive ability, investment analysts can use it to assess a chain's intangible value, given a set of characteristics.

At an operational level, the study's findings generate action points for managers of firms (in the four segments) to maximize their firms' intangible values:

Segment 1. For the poorly performing, medium-sized firms in Segment 1, a reduced emphasis on market-based channels, especially at moderate levels and when they have high financial liquidity, and an increased scope of vertical integration will reduce the negative effects of dual distribution on their intangible values. Firms in Segment 1, with high financial liquidity, will be well served from a shareholder value perspective by investing their cash resources in vertically integrated channels. Alternatively, these firms can invest their cash resources in additional advertising because as their advertising stocks increase, an increased emphasis on market-based channels increases their intangible value.

Segment 2. The large firms with large advertising stocks in Segment 2 can increase their intangible values by emphasizing market-based channels because dual distribution has both positive main and interaction effects with the scope of their vertical integration and advertising stocks. The negative effect of the interaction between dual distribution and financial leverage suggests that these firms will be well served from a shareholder value perspective by reducing their financial leverage.

Segment 3. In general, the shareholder values of the small firms with a large portfolio of brands in Segment 3 appear to be impervious to their dual distribution strategies. One exception is the negative effect of the interaction between their dual distribution and financial leverage on intangible values. These firms can increase their intangible values by decreasing their financial leverage.

Segment 4. Surprisingly, for the very small, well-performing firms, there are no positive effects of dual distribution, either independently or jointly with other characteristics. However, dual distribution has a negative interaction effect with age, advertising, and financial liquidity, suggesting that over time and as their advertising stocks and liquidity increase, the stock market rewards them for investing in vertically integrated channels.

Finally, a study of performance in franchise firms is worthy of empirical investigation in its own right because franchising is a vital force in the U.S. economy (Dant and Kaufmann 2003). In 2003, business format franchises accounted for more than 600,000 establishments, 8 million jobs, and $460 billion of economic output (PricewaterhouseCoopers 2004). Restaurant chains dominate the industry, and 15 of the 20 largest franchises are restaurants with sales representing 54% of revenues of all U.S. franchises. Yet there are few insights into the performance of franchise firms. As Bradach (1997, p. 301) notes, "future work should investigate how the functioning and performance of the plural forms compares in chains with different mixes of company and franchised units." The current study's findings on the factors affecting a chain's performance could be useful to managers of franchise firms in general and to managers of restaurant chains in particular to increase their firm value.

Limitations and Opportunities for Further Research

In this first study on the performance implications of dual distribution, I focused on franchising in general and restaurant chains in particular. The choice of one industry allows for the control of cross-industry variation, enabling a clean test of the hypotheses. However, as a result, I do not consider environmental characteristics (e.g., uncertainty, competitive conditions), which have been shown to be pertinent to channel decisions. In addition, the use of secondary data precluded consideration of organizational factors (e.g., culture, trust, organizational structure) that affect channel management. Further research on the performance implications of dual distribution that uses other methods (e.g., in-depth interviews, surveys) across diverse industry contexts incorporating other environmental and organizational characteristics would be useful.

This article's focus on dual distribution raises questions about the generalizability of the study's findings to other plural governance settings. The study's sample (N = 55), though representative of the population of restaurant chains, is small. Further research on performance in other plural governance settings (e.g., foreign market entry, outsourcing) with larger sample sizes would represent useful extensions.

The dependent variable in this study is intangible firm value, as measured by Tobin's q. Although Tobin's q is a reliable measure of firm value, it represents the stock market's evaluation of the level, speed, and risk of the firm's future cash flows. Research extensions using alternative performance metrics, including survival, risk, and volatility in earnings, would both complement and extend the study's findings.

In summary, this study represents a useful first step in exploring the performance implications of dual distribution strategies. I hope that this article stimulates further work that relates plural governance in general and dual distribution in particular to firm performance.

Appendix

Outline of Latent Class Regression Models

Let f = 1, ..., F index the restaurant chains; t = 1, ..., T_f index the year of restaurant chain f; s = 1, ..., S index the number of latent segments; Y_{ft} be the Tobin's q of restaurant chain f in year t; and X_{ftp} be the value of the pth explanatory variable for restaurant chain f in year t. Assume that the vector of observations for Y_{ft} arises from a population that is a mixture of S segments in proportions π_1, π_2, π_3, π_s, where a priori membership of the observations in the different segments is unknown. Then, each firm belongs to only one segment, and the point mass for each segment is denoted by π_s, such that

$$\text{(A1)} \qquad \sum_{s=1}^{S} \pi_s = 1, \text{ and } 0 \leq \pi_s \leq 1,$$

where π_s is the point probability of a firm belonging to segment s.

The objective is to predict the means of the dependent variable for each observation in each segment using a set of explanatory variables. To that end, I specify a linear predictor η_{fts}, which is a function of the P explanatory variables, $X_1, X_2, ..., X_p$ ($X_p = X_{ftp}$; p = 1, ..., P), and parameter vectors $\beta = \beta_{sp}$ to be estimated in segment s:

(A2) $$\eta_{fts} = \sum_{p=1}^{P} X_{ftp}\beta_{sp}.$$

In turn, the linear predictor is related to the mean of the distribution, μ_{st}, through a link function $g(\cdot)$ such that in segment s,

(A3) $$\eta_{fts} = g(\mu_{fts}).$$

In the current case, given the panel structure of the data, the Y_{ft} values are not independent across the t time periods. The distribution of Y_{ft}, given that it comes from segment s, $f_s(Y_{ft}|\theta_s)$ is normal within each of the s segments, so that the identity link function is pertinent such that $\eta_{fts} = \mu_{st}$. Thus, by combining Equations A2 and A3, I obtain the standard linear regression model within the segments, given the assumption that the firm can belong to only one segment. I model π_s with a logit formulation. Specifically,

(A4) $$\pi_s = \frac{e^{\delta_s}}{\sum_{s=1}^{S} e^{\delta_s}},$$

where δ_s is the segment-specific vector of profiling coefficients. Then, Equation A4 is standardized as

(A5) $$\pi_s = \frac{e^{\delta_s}}{1 + \sum_{s=1}^{S-1} e^{\delta_s}},$$

and δ_s is parameterized as

(A6) $$\delta_s = \beta_{0s} + \sum_{p=1}^{P} z_{fp}\beta_{ps},$$

where z_{fp} is the matrix of profiling variables; β_{0s} is a segment-specific constant term; β_{ps} is a vector of coefficients for the profiling variables $\forall$ s = 1, …, S – 1; and P is the number of profiling variables.

Estimation and Model Selection

If the distributions conditional on the segments are formulated, the unconditional distribution of Y_{ft} is obtained as follows:

(A7) $$f(Y_{ft}|\varphi) = \sum_{s=1}^{S} \pi_s f_s(Y_{ft}|\theta_s),$$

where $\varphi = (\pi, \theta)$ and $\theta = (\beta_f, \lambda_s)$. The purpose of the latent class regression estimation is to estimate the parameter vector $\varphi = (\pi, \theta)$ using the maximum likelihood method, so that the likelihood for φ is

(A8) $$L(\varphi; Y) = \prod_{f=1}^{F}\prod_{t=1}^{T} f(Y_{ft}|\varphi).$$

For a prespecified value of the number of segments s, an estimate of φ (φ_0) can be obtained by maximizing the likelihood Equation A8 with respect to φ, subject to the restrictions in Equation A1, so that the parameter vector φ could have produced the observed vector Y. To obtain the parameter estimates, I use the expectation-maximization algorithm to maximize the natural logarithm of the likelihood function specified in Equation A8. I use 50 randomly selected starting values for each of the models to ensure convergence (Wedel and Kamakura 2000). To obtain the standard errors for the parameters, the Hessian (or the information matrix) for the final parameter estimates is inverted. After an estimate of φ has been obtained, estimates of the posterior probability p_{fs}, a probabilistic allocation of the firm to the latent segments, are obtained using Bayes' theorem as follows:

(A9) $$p_{fs} = \frac{\pi_s f_s(y_f|\theta_s)}{\sum_{s=1}^{S} \pi_s f_s(y_f|\theta_s)}.$$

To determine the number of latent segments, the BIC (Schwarz 1978) is used to compare the model with s segments with a model with s + 1 segments $\forall$ s = 1, 2, …, until the model fit stops improving. The BIC is calculated as BIC = –2 × LL + P × ln(F), where LL, P, and F stand for log-likelihood value, number of parameters, and sample size, respectively. I also report an entropy measure of separation (ES), or the extent of separation of the segments, which is calculated as

(A10) $$ES = 1 - \frac{\sum_{f=1}^{F}\sum_{s=1}^{S} p_{fs}\ln(p_{fs})}{F\ln(S)},$$

where p_{fs} is the probability of firm f belonging to segment s, which is calculated using the Bayes rule specified in Equation A9. The ES is bounded in the range 0–1, such that a value closer to 1 indicates good separation of the latent segments.

REFERENCES

Aaker, David A. and Robert Jacobson (2001), "The Value Relevance of Brand Attitude in High-Technology Markets," *Journal of Marketing Research*, 38 (November), 485–93.

Alchian, A.A. and H. Demsetz (1972), "Production Information Costs and Economic Organization," *American Economic Review*, 62 (December), 777–95.

Amir, Eli and Baruch Lev (1996), "Value Relevance of Non-Financial Information: The Wireless Communications Industry," *Journal of Accounting and Economics*, 22 (1–3), 3–30.

Anderson, Erin (1988), "Transactions Costs as Determinants of Opportunism in Integrated and Independent Sales Forces," *Journal of Economic Behavior and Organization*, 9 (May), 247–64.

———, George S. Day, and V.K. Rangan (1997), "Strategic Channel Design," *Sloan Management Review*, 38 (4), 59–69.

——— and David Schmittlein (1984), "Integration of the Sales Force: An Empirical Investigation," *RAND Journal of Economics*, 15 (3), 385–95.

——— and Barton A. Weitz (1986), "Make or Buy Decisions: Vertical Integration and Marketing Productivity," *Sloan Management Review*, 27 (Spring), 3–20.

Anderson, Eugene W., Claes Fornell, and Sanal K. Mazvancheryl (2004), "Customer Satisfaction and Shareholder Value," *Journal of Marketing*, 68 (October), 172–85.

Arrow, Kenneth (1985), "The Economics of Agency" in *Principals and Agents: The Structure of Business*, J. Pratt and R. Zeckhauser, eds. Boston: Harvard Business School Press.

Banerji, Shumeet and Carol Simon (1991), "Franchising Versus Ownership: A Contracting Explanation," mimeo, Graduate School of Business, University of Chicago.

Barth, Mary E., Michael B. Clement, George Foster, and Ron Kasznik (1998), "Brand Values and Capital Market Evaluation," *Review of Accounting Studies*, 3 (1–2), 41–68.

Bayus, Barry L., Gary Erickson, and Robert Jacobson (2003), "The Financial Rewards of New Product Introductions in the Personal Computer Industry," *Management Science*, 49 (2), 197–209.

Belsley D., E. Kuh, and R. Welsch (1980), *Regression Diagnostics*. New York: John Wiley & Sons.

Bergen, Mark, Shantanu Dutta, and Orville C. Walker Jr. (1992), "Agency Relationships in Marketing: A Review of the Implications and Applications of Agency and Related Theories," *Journal of Marketing*, 56 (July), 1–24.

Bharadwaj, Anandhi S., Sundar G. Bharadwaj, and Benn R. Konsynski (1999), "Information Technology Effects on Firm Performance as Measured by Tobin's Q," *Management Science*, 45 (June), 1008–1024.

Blair, R.D. and D.L. Kaserman (1994), "A Note on Incentive Compatibility Under Franchising," *Review of Industrial Organization*," 9 (3), 323–30.

Bolton, Ruth (2004), "Linking Marketing to Financial Performance and Firm Value," *Journal of Marketing*, 68 (October), 73–75.

Bond, Stephen and Costas Meghir (1994), "Dynamic Investment Models and the Firm's Financial Policy," *Review of Economic Studies*, 61 (207), 197–222.

Bowman, E.H. and Hurry, D. (1993), "Strategy Through the Option Lens: An Integrated View of Resource Investments and the Incremental-Choice Process," *Academy of Management Review*, 18 (4), 760–82.

Bradach, Jeffrey F. (1997), "Using the Plural Form in the Management of Restaurant Chains," *Administrative Science Quarterly*, 42 (2), 276–303.

——— (1998), *Franchise Organizations*. Boston: Harvard Business School Press.

Bradley, M., G.A. Jarrell, and E.H. Kim (1984), "On the Existence of an Optimal Capital Structure," *Journal of Finance*, 39 (3), 857–78.

Brickley J.A. and F.H. Dark (1987), "The Choice of Organizational Form: The Case of Franchising," *Journal of Financial Economics*, 18 (2), 401–420.

Brüderl, Josef and Rudolf Schussler (1990), "Organizational Mortality: The Liabilities of Newness and Adolescence," *Administrative Science Quarterly*, 35 (3), 530–47.

BusinessWeek (1997), "McDonald's Aggressive Growth Strategy Isn't Going Down Well with Rebel Franchisees," (June 2), 34–36.

Capon, Noel, John U. Farley, and S. Hoenig (1990), "Determinants of Financial Performance: A Meta-Analysis," *Management Science*, 36 (October), 1143–59.

Carney, M. and E. Gedajlovic (1991), "Vertical Integration in Franchise Systems: Agency Theory and Resource Explanations," *Strategic Management Journal*, 12 (8), 607–629.

Chung, Kee H. and Stephen W. Pruitt (1994), "A Simple Approximation of Tobin's Q," *Financial Management*, 23 (Autumn), 70–74.

Comanor, W.S. and T.A. Wilson (1967), "Advertising, Market Structure and Performance," *Review of Economics and Statistics*, 49 (4), 423–40.

Dahlstrom, Robert and Arne Nygaard (1994), "A Preliminary Investigation of Franchised Oil Distribution in Norway," *Journal of Retailing*, 70 (2), 179–91.

——— and ——— (1999), "An Empirical Investigation of Ex Post Transaction Costs in Franchised Distribution Channels," *Journal of Marketing Research*, 36 (May), 160–70.

Dant, Rajiv and Patrick Kaufmann (2003), "Structural and Strategic Dynamics in Franchising," *Journal of Retailing*, 79 (2), 63–75.

Dutta, Shantanu, Mark Bergen, Jan B. Heide, and George John (1995), "Understanding Dual Distribution: The Case of Reps and House Accounts," *Journal of Law, Economics, and Organization*, 11 (1), 189–204.

Fama, Eugene F. and Michael C. Jensen (1983), "Agency Problems and Residual Claims," *Journal of Law and Economics*, 26 (2), 327–49.

Fazzari, S.M., R.G. Hubbard, and B. Petersen (1988), "Financing Constraints and Corporate Investment," *Brookings Papers on Economic Activity*, (Spring), 141–206.

Gatignon, Hubert and Dominique M. Hanssens (1987), "Modeling Marketing Interactions with Application to Sales Force Effectiveness," *Journal of Marketing Research*, 24 (August), 247–57.

Geyskens, Inge, Katrijn Gielens, and Marnik G. Dekimpe (2002), "The Market Valuation of Internet Channel Additions," *Journal of Marketing*, 66 (April), 102–119.

Grossman, S.J. and O.D. Hart (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94 (4), 691–719.

Hambrick, D.C. (1983), "An Empirical Typology of Mature Industrial Product Environments," *Academy of Management Journal*, 26 (2), 213–30.

Heckman, James J. (1979), "Sample Selection Bias as a Specification Error," *Econometrica*, 47 (January), 153–62.

Heide, Jan B. (1994), "Interorganizational Governance in Marketing Channels," *Journal of Marketing*, 58 (January), 71–85.

——— (2003), "Plural Governance in Industrial Purchasing," *Journal of Marketing*, 67 (October), 18–29.

Hirschey, Mark and Jerry Weygandt (1985), "Amortization Policy for Advertising and Research and Development Expenditures," *Journal of Accounting Research*, 23 (1), 326–35.

Hsiao, Cheng and A. Kamil Tahmiscioglu (1997), "A Panel Analysis of Liquidity Constraints and Firm Investment," *Journal of the American Statistical Association*, 92 (438), 455–65.

Jensen, M. and W. Meckling (1976), "Theory of the Firm: Managerial Behavior: Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3 (4), 305–360.

John, George and Barton A. Weitz (1988), "Forward Integration into Distribution: An Empirical Test of Transaction Cost Analysis," *Journal of Law, Economics, and Organization*, 4 (2), 337–55.

Kulatilaka, Nalin and L. Trigeorgis (1994), "The General Flexibility to Switch: Real Options Revisited," *International Journal of Finance*, 6 (2), 778–98.

Lafontaine, Francine (1992), "Agency Theory and Franchising: Some Empirical Results," *RAND Journal of Economics*, 23 (2), 263–83.

——— and Patrick J. Kaufmann (1994), "The Evolution of Ownership Patterns in Franchise Systems," *Journal of Retailing*, 70 (2), 97–113.

——— and Kathryn L. Shaw (2005), "Targeting Managerial Control: Evidence from Franchising," *RAND Journal of Economics*, 36 (1), 131–50.

Lang, Larry, Eli Ofek, and R. Stulz (1996), "Leverage, Investment, and Firm Growth," *Journal of Financial Economics*, 40 (1), 3–30.

Lee, Ruby P. and Rajdeep Grewal (2004), "Strategic Responses to New Technologies and Their Impact on Firm Performance," *Journal of Marketing*, 68 (October), 157–71.

Lewin-Solomons, S. (1998), "The Plural Form in Franchising: A Synergism of Market and Hierarchy," working paper, Department of Economics, Iowa State University.

March, James G. (1991), "Exploration and Exploitation in Organizational Learning," *Organization Science*, 2 (1), 71–87.

Martin, R.E. (1988), "Franchising and Risk Management," *American Economic Review*, 78 (5), 954–68.

Mathewson, George and Ralph Winter (1984), "An Economic Theory of Vertical Restraints," *RAND Journal of Economics*, 15 (Spring), 27–38.

McDonald's Corporation (1995), "McDonald's 1995 Annual Report," (accessed June 22, 2005), [available at http://www.edgaronline.org].

Mizik, Natalie and Robert Jacobson (2003), "Trading Off Between Value Creation and Value Appropriation: The Financial Implications of Shifts in Strategic Emphasis," *Journal of Marketing*, 67 (January), 63–76.

Moriarty, Rowland T. and Ursula Moran (1990), "Managing Hybrid Marketing Systems," *Harvard Business Review*, 68 (6), 146–55.

Myers, Stewart C. (1977), "Determinants of Corporate Borrowing," *Journal of Financial Economics*, 5 (2), 147–76.

——— and N. Majluf (1984), "Corporate Finance and Investment Decisions When Firms Have Information Investors Do Not Have," *Journal of Financial Economics*, 13 (2), 187–221.

——— and Raghuram G. Rajan (1998), "The Paradox of Liquidity," *The Quarterly Journal of Economics*, 113 (3), 733–71.

Oxenfeldt, A. and A.O. Kelly (1969), "Will Successful Franchise Systems Ultimately Become Wholly-Owned Chains?" *Journal of Retailing*, 44 (4), 69–83.

Papa John's International (2004), "Papa John's Annual Report," (accessed June 27, 2005), [available at http://www.papajohns.com].

Pauwels, Koen, Jorge Silva-Risso, Shuba Srinivasan, and Dominique M. Hanssens (2004), "New Products, Sales Promotions, and Firm Value: The Case of the Automobile Industry," *Journal of Marketing*, 68 (October), 142–56.

Pedhazur, Elazar J. (1997), *Multiple Regression in Behavioral Research*. New York: Harcourt Brace College Publishers.

Penrose, Edith T. (1959), *The Theory of the Growth of the Firm*. New York: John Wiley & Sons.

PricewaterhouseCoopers (2004), "The Economic Impact of Franchised Businesses" a study for the International Franchise Association Educational Foundation, (accessed April 7, 2004), [available at http://www.franchise.org/files/EIS6_1.pdf].

Rao, Vithala R., Manoj Agarwal, and Denise Dahloff (2004), "How Is Manifested Branding Strategy Related to the Intangible Value of a Corporation?" *Journal of Marketing*, 68 (October), 126–41.

Ravenscraft D. and F.M. Scherer (1982), "The Lag Structure of Returns to Research and Development," *Applied Economics*, 14 (6), 603–620.

Rindfleisch, Aric and Jan B. Heide (1997), "Transaction Cost Analysis: Past, Present, and Future Applications," *Journal of Marketing*, 61 (October), 30–54.

Rust, Roland T., Tim Ambler, Gregory S. Carpenter, V. Kumar, and Rajendra K. Srivastava (2004), "Measuring Marketing Productivity: Current Knowledge and Future Directions," *Journal of Marketing*, 68 (October), 76–89.

Schmidgall, Raymond S., David K. Hayes, and Jack D. Ninemeier (2002), *Restaurant Financial Basics*. Hoboken, NJ: John Wiley & Sons.

Schwarz, Gideon (1978), "Estimating the Dimensions of a Model," *Annals of Statistics*, 6 (2), 461–64.

Simon, Carol J. and Mary W. Sullivan (1993), "The Measurement and Determinants of Brand Equity: A Financial Approach," *Marketing Science*, 12 (Winter), 28–52.

Srivastava, Rajendra K., Tassaduq A. Shervani, and Liam Fahey (1998), "Market-Based Assets and Shareholder Value: A Framework for Analysis," *Journal of Marketing*, 62 (January), 2–18.

———, ———, and ——— (1999), "Marketing, Business Processes, and Shareholder Value: An Organizationally Embedded View of Marketing Activities and the Discipline of Marketing," *Journal of Marketing*, 63 (Special Issue), 168–79.

Stern, Louis W. and Barton A. Weitz (1997), "The Revolution in Distribution: Challenges and Opportunities," *Long Range Planning*, 30 (6), 823–29.

Wedel, M. and Wagner A. Kamakura (2000), *Market Segmentation: Conceptual and Methodological Foundations*. Boston: Kluwer Academic.

Weiss, Allen M. and Erin Anderson (1992), "Converting from Independent to Employee Sales Forces: The Role of Perceived Switching Costs," *Journal of Marketing Research*, 22 (February), 101–115.

Williamson, Oliver (1975), *Markets and Hierarchies*. New York: The Free Press.

——— (1985), *The Mechanisms of Market Governance*. New York: Oxford University Press.

——— (1991), "Comparative Economic Organization: The Analysis of Discrete Structural Alternatives," *Administrative Science Quarterly*, 36 (2), 269–96.

Yin, Xiaoli and Edward J. Zajac (2004), "The Strategy/Governance Structure Fit Relationship: Theory and Evidence in Franchising Arrangements," *Strategic Management Journal*, 25 (4), 365–83.

Zeithaml, Valarie A., P. Rajan Varadarajan, and Carl P. Zeithaml (1988), "The Contingency Approach: Its Foundations and Relevance to Theory Building and Research in Marketing," *European Journal of Marketing*, 22 (7), 37–64.

Jan-Benedict E.M. Steenkamp & Inge Geyskens

How Country Characteristics Affect the Perceived Value of Web Sites

The authors examine how country characteristics systematically moderate the effects of individual-level drivers of the perceived value that consumers derive from visiting a brand manufacturer's Web site. They test hypotheses on data collected from 8886 consumers from 23 countries on three continents, involving 30 Web sites of the world's largest consumer packaged goods companies. They find that the effect of privacy/security protection on perceived value is stronger for people from countries with a weak rule of law, whereas people from countries that are high on national identity give more weight to whether there is cultural congruity between the site and themselves. People who live in more individualistic countries give more weight to pleasure, to privacy/security protection, and to customization in their perceived value judgments than people from collectivistic countries. The authors discuss implications for Web site design strategies.

The emergence and astonishing growth of the Internet is easily one of the most important developments in societies around the globe. Increasingly, companies are creating Web sites in an attempt to add value for their consumers. Both academics and practitioners have emphasized the key role of consumer value delivery in company success (Day 1990; Gale 1994; Naumann 1995; Woodruff 1997; Zeithaml 1988), including the success of Internet-based activities (Zeithaml, Parasuraman, and Malhotra 2002).

Given the global nature and reach of the Internet, consumers across the world can access Web sites. Although academics and managers agree about which factors (e.g., privacy/security, customization) affect consumer value, there is a lack of research into how the effectiveness of various factors in creating consumer value may vary across countries. The primary purpose of this study is to test whether relationships between consumer value and its drivers are systematically and predictably moderated by the institutional context in which consumers live.

We investigate this issue in the context of Web sites operated by large consumer packaged goods (CPG) companies. Whereas Web sites may be designed for various commercial uses, including being a communication channel and offering an (additional) online sales channel, the focus of this article is on CPG sites that are primarily designed as a communication tool to forge cognitive and emotional bonds between a company's customers and its brands. Since consumer uptake of the Internet has rocketed, the Internet has become a major avenue for CPG companies to reach their consumers. In July 2003, the top ten sites of food and beverage makers attracted 9.5 million unique visitors, more than 75% more than a year before and only slightly fewer, for comparison's sake, than the number of visitors to the top ten university sites (Tedeschi 2003). Yet, in bonding with their consumers, "more [CPG] companies have failed than succeeded" (Curry 1999, p. 20), and 15% of the top CPG brands still lack a Web site because of a lack of understanding of what their target consumers value in using the Internet (*New Media Age* 2003).

The contribution of this study is twofold. First, we develop a conceptual model that pertains to the moderating role of the institutional context in which consumers live on the effects of individual and Web site characteristics. Second, we test the hypotheses using a large-scale data set, involving 30 Web sites of 16 of the world's largest CPG companies from 8 different European countries and more than 8500 visitors from 23 countries on three continents (Australasia, Europe, and North America).

In the following sections, we introduce the main effect hypotheses. Given previous support in the literature for these relationships, this introduction is brief and primarily serves as a basis to develop the focal country-level moderating hypotheses. These sections are followed by a discussion of the methods we used to test the hypotheses. Finally, we provide the results and discuss the implications of our findings.

Drivers of Perceived Value

The focal construct in our study is the "perceived value" that consumers derive from visiting a brand manufacturer's Web site. Drawing on Holbrook's definition and theory of

---

Jan-Benedict E.M. Steenkamp is C. Knox Massey Professor of Marketing and Marketing Area Chair, Kenan-Flagler Business School, University of North Carolina, Chapel Hill (e-mail: JBS@unc.edu). Inge Geyskens is Associate Professor of Marketing, Tilburg University, the Netherlands (e-mail: I.Geyskens@uvt.nl). The authors gratefully acknowledge AiMark, the European Brands Association (AIM), MetrixLab, and the participating packaged goods companies (Barilla, Beiersdorf, Cadbury Schweppes, Colgate-Palmolive, Danone, Heineken, Heinz, Henkel, Kraft Foods, Lindt & Sprüngli, Mars, Procter & Gamble, Reckitt Benckiser, Sara Lee, SCA Hygiene, and Unilever) for their help in developing the questionnaire and for their assistance in collecting the data. They also thank the anonymous *JM* reviewers, Marnik Dekimpe, and Els Gijsbrechts for valuable comments and suggestions. Finally, they extend their appreciation to the Netherlands Organization for Scientific Research for financial assistance.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

perceived value (Holbrook 1994, 1999; Holbrook and Corfman 1985), we define the perceived value of a Web site as an interactive, relativistic, preference experience that results from visiting the Web site. It is "interactive" in that a value judgment is neither wholly subjective nor wholly objective but rather entails a subject–object interaction, meaning that it is affected by aspects of the person (e.g., emotions, sociodemographics) visiting the Web site and the characteristics of the Web site (e.g., privacy/security). It is "relativistic" in that value judgments differ among Web site visitors, just as consumer characteristics and perceptions of Web site characteristics may differ among people. Value involves "preference," which Holbrook and Corfman (1985, p. 40) interpret as "favorable disposition"; this encompasses evaluations (e.g., the Web site is useful) and intentions to act on these evaluations (e.g., intention to visit the Web site again; see also Chen, Clifford, and Wells 2002; Wells and Prensky 1996). Finally, value is an "experience" in that it resides in the consumption of the object, that is, in surfing the Web site.

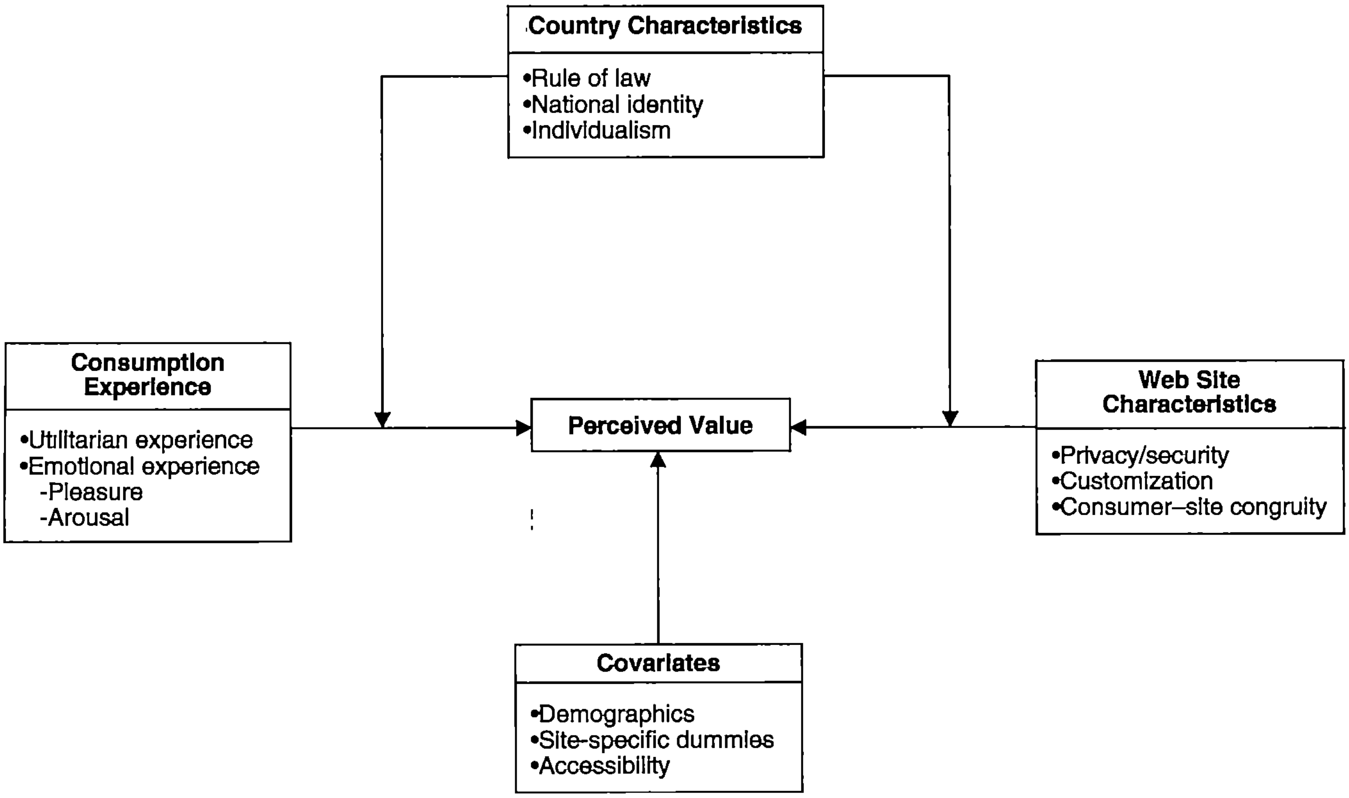
Figure 1 presents our conceptual framework, delineating drivers of perceived value of Web sites. We discuss the hypotheses underlying this framework subsequently.

Consumption Experience

The consumption experience has both a cognitive (utilitarian) and an emotional component (Holbrook and Hirschman 1982). This also applies to the Internet context; thus, it becomes crucial to evaluate "Web sites in terms of the extent to which they deliver these two types of experience" (Novak, Hoffman, and Yung 2000, p. 39). By visiting a Web site, consumers may acquire useful information to help them in their decision process. That is, the Web site may be instrumental in the attainment of utilitarian consumption goals, which is valued by consumers (Shankar, Smith, and Rangaswamy 2003).

Consumers may also derive value from the emotions generated during the Web site visit per se. In many cases, the consumer can even achieve flow while visiting the site (Novak, Hoffman, and Yung 2000). Various emotion typologies have been proposed in the literature, of which the pleasure–arousal typology is among the most prominent (e.g., Eliashberg and Sawhney 1994; Olney, Holbrook, and Batra 1991; Reisenzein 1994). "Pleasure" refers to the affective feelings generated by the Web site, and "arousal" pertains to the level of activation or excitement generated by the Web site. The higher the feelings of pleasure evoked by a site, the more positive its evaluation will be. On the basis of the Wundt curve (Berlyne 1971), we hypothesize that the relationship between the arousal evoked in a person by visiting the Web site and that person's evaluation of the Web site takes the form of an inverted U shape; that is, moderate degrees of arousal are preferred over low and high levels of arousal. Formally,

FIGURE 1
Conceptual Framework



> H_1: The perceived value derived from a Web site visit (a) is positively related to the utilitarian experience associated with visiting the site, (b) is positively related to the pleasure associated with visiting the site, and (c) exhibits an inverted U-shaped relationship with the arousal induced by the site.

Web Site Characteristics

Two Web site characteristics that are consistently mentioned in the literature as being among the most important characteristics affecting consumer evaluations of sites are privacy/security and customization (Sultan et al. 2002; Zeithaml, Parasuraman, and Malhotra 2002). Given the international scope of our study, a third characteristic appears pertinent: the cultural congruity between the site and its visitor (Luna, Peracchio, and De Juan 2002).

Perceived privacy and security are two related key drivers of consumer value (Hoffman, Novak, and Peralta 1999; Zeithaml, Parasuraman, and Malhotra 2002). Both involve protecting Web site users' personal information. The higher a site scores on privacy and security protection in the minds of consumers, the more the risk that consumers perceive to exist when supplying private information online will be mitigated (Lwin and Williams 2004), and thus perceived value will increase.

The Internet also provides unique opportunities to customize information. Customization refers to the extent to which the Web site's content is adapted to an individual's unique needs versus being presented as a standardized package. People value personalized experiences (see, e.g., Feldman and Lynch 1988).

The Internet allows global access to a Web site. This offers the opportunity for cross-border information flow, but it also means that there may be cultural incongruity between the site and the visitor. Cultural congruity refers to whether the Web site is translated into the visitor's local country-specific language; offers local content; and exhibits localization in the form of country-specific dates, zip codes, currency formats, and so forth (Singh, Furrer, and Ostinelli 2004). Exposure to a culturally congruous site facilitates processing of the site's content (Luna, Peracchio, and De Juan 2002) and conveys the marketer's respect for the visitor's culture (Koslow, Shamdasani, and Touchstone 1994), both of which consumers value. Thus:

> H_2: The perceived value derived from a Web site visit (a) is positively related to the perceived privacy/security protection of the site, (b) is positively related to the site's degree of customization, and (c) is higher when there is cultural congruity between the site and the consumer.

The Moderating Role of a Country's Institutional Context

Web site visitors are members of larger institutional systems and thus are subject to the conflicts and compatibilities between their own evaluative structure and the priorities of the institutional system. Institutions are "the most enduring features of social life ... giving 'solidity' [to social systems] across time and space" (Giddens 1984, p. 24). A society's institutional priorities reflect the basic issues and problems that societies must confront to regulate human activity. The shared institutional priorities in a society help shape the social and economic reward contingencies to which people must adapt to function smoothly and effectively (Schwartz 1994).

A society's institutional priorities stimulate the activation of beliefs that are in line with and conducive to maintaining these priorities, and beliefs that run counter to these priorities are discouraged (Scott 2001; Triandis 1989). Thus, as standards, these priorities influence how individual beliefs and emotions are evaluated and what kinds of expressions are encouraged or discouraged. These positive or negative institutional reinforcement mechanisms, which operate between the two types of constructs, give rise to interactions between society-level and individual-level variables (Schwartz 1994).

Scott (2001) distinguishes three interrelated but distinct "pillars of institutions": the regulative, moral, and cultural systems (see also Grewal and Dharwadkar 2002; Suchman 1995). Each of these systems has been identified by various social theorists as a vital component of a country's institutional context (Scott 2001). These three systems form a continuum that moves "from the conscious to the unconscious, from the legally enforced to the taken for granted" (Hoffman 2001, p. 36). We summarize some of the key characteristics of the three pillars of institutions in Table 1.

The "regulative" system involves the capacity to establish formal rules, inspect society members' conformity to them, and, if necessary, impose sanctions (Scott 2001). It includes the presence and efficacy of regulatory institutions and the associated legal system that exists to ensure stability, order, and continuity of societies (Arndt 1979). The "moral" system involves societal norms of right and wrong and "introduces a prescriptive, evaluative, and obligatory dimension in social life" (Scott 2001, p. 54). It defines goals or objectives (e.g., making a profit, loving one's country) but also specifies appropriate ways to pursue them (e.g., conceptions of fair business rules, purchase of domestic products). The "cultural" system represents the culturally supported beliefs, attitudes, habits, and behaviors. It maintains that external cultural frameworks shape internal interpretive processes and shared understanding (Triandis 1989). Certain routines (e.g., shaking hands when meeting someone) are followed because they are taken for granted "as the way we do these things" (Scott 2001, p. 57). In this study, we include one key component of each of these pillars of institutions, namely, rule of law (regulative; Oxley and Yeung 2001), national identity (moral; Giddens 1981), and national-cultural individualism (cultural; Oyserman, Coon, and Kemmelmeier 2002).

Regulative System: Rule of Law

In general, it is recognized that the rule of law is an essential component of the regulative subsystem (Licht, Goldschmidt, and Schwartz 2003; North 1990; Oxley and Yeung 2001; Scott 2001). It is a central tenet in the "good governance" agenda of the major international institutions, such as the International Monetary Fund and the World Bank (Kaufmann, Kraay, and Zoido-Lobatón 1999). These bodies define governance as the traditions and institutions by

TABLE 1
Three Pillars of Institutions

	Regulative	Moral	Cultural
Basis of compliance	Expedience	Social obligation	Taken-for-grantedness Shared understanding
Basis of order	Regulative rules	Binding expectations	Constitutive schema
Mechanisms	Coercive	Normative	Mimetic
Basis of legitimacy	Legally sanctioned	Morally governed	Comprehensible Recognizable Culturally supported
Primary scientific domain	Economics	Sociology Psychology	Anthropology Sociology

Source: Adapted from Scott (2001).

which authority in a country is exercised for the common good. The rule of law addresses the degree to which the behavior of individuals and organizations (including government authorities) is guided by formal, transparent, legal rules that apply to everyone, regardless of rank and status (Licht, Goldschmidt, and Schwartz 2003). A country with a strong rule of law has sound political institutions and a strong, independent court system. Its citizens are entitled to equal protection under the law and are willing to accept the established institutions to make and implement laws and resolve disputes (Oxley and Yeung 2001).

A strong rule of law generates greater transparency and stability regarding the boundaries of acceptable behavior. It reduces Web site visitors' uncertainty about what legal protection they can expect and enhances their ability to successfully litigate at least the most serious cases of fraudulent online dealings. In countries in which the rule of law is weak, this ability is undermined and online visitors have greater need for other safeguards against fraud and misuse of their trust. A particularly important type of online safeguard is the privacy/security protection offered by the Web site (Zeithaml, Parasuraman, and Malhotra 2002), which therefore is expected to be (even) more important in countries that are characterized by a weak rule of law.

> H_3: The effect of perceived privacy/security protection on perceived value is stronger in countries with a weak rule of law.

Moral System: National Identity

Previous research in marketing has noted the important role of one particular moral construct in explaining consumer attitudes and behavior: consumer ethnocentrism (Balabanis et al. 2001; Shimp and Sharma 1987; Steenkamp, Batra, and Alden 2003), defined as "the beliefs held by [American] consumers about the appropriateness, indeed *morality* of buying foreign-made products" (Shimp and Sharma 1987, p. 280, emphasis added). There are systematic differences in consumer ethnocentrism among countries (e.g., Balabanis et al. 2001).

The construct of national identity is conceptually closely related to consumer ethnocentrism (cf. Baughn and Yaprak 1996) in that both constructs provide prescriptions for morally motivated behavior. However, whereas consumer ethnocentrism is a consumer-domain-specific construct, national identity is a more abstract, "central" construct (cf. Lastovicka 1982) and, as such, is more central to a society's moral system (Scott 2001). The theoretical basis of national identity is social identity theory, which distinguishes between in-groups and out-groups and posits the incorporation of the in-group as part of the self of its members (Mackie and Smith 1998; Tajfel 1981). Various theorists have argued that nations can be viewed as communities, whose members form a social group. For example, Giddens (1981) stated that a nation's inhabitants have "an overall awareness ... of belonging to an inclusive community with a certain identity." In a similar vein, Druckman (1994, p. 63) argues that nations "achieve personal relevance for individuals when they ... gain a sense of identity and self-esteem through their national identification."

Countries that are high on national identity take pride in and display a positive bias toward their country's symbols, language, and culture (Billig 1995). These daily reminders of national identity, as well as its omnipresence in education and mass media, provide a continuous reinforcement of the feeling of belonging to a national group (Verlegh 2001). The level of national identification differs widely among countries (International Social Survey Program [ISSP] 1998), and differences in national identity are related to differential tendencies to favor culturally adapted products and other culturally congruous symbols (cf. Kowalski and Wolfe 1994; Steenkamp, Batra, and Alden 2003). Consistent with this body of research, we expect that people from countries that are high on national identity attach more importance to the cultural congruity of the Web site in shaping their value perceptions.

> H_4: The effect of cultural congruity on perceived value is stronger in countries characterized by high national identity.

Cultural System: National-Cultural Individualism

National-cultural individualism pertains to the degree to which people in a country prefer to act as individuals. Conversely, collectivism pertains to the extent to which people in a country prefer to act as members of a group. People in individualistic societies place their personal goals, motivations, and desires ahead of those of others, whereas collectivistic cultures are conformity oriented and show a higher degree of group behavior and concern to promote their continued existence (Kagitcibasi 1997). Individualism's "pivotal role" (Allik and Realo 2004, p. 32) in the cultural system has been consistently emphasized by social science researchers (Oyserman, Coon, and Kemmelmeier 2002). It has been studied as such in a wide range of disciplines, including philosophy, history, anthropology, sociology, psychology, and business (Triandis 1995). Kagitcibasi (1997, p. 3) notes that "about one-third of recently published studies cited this construct [individualism/collectivism] as at least a partial explanation of observed cross-cultural differences."

Psychologists have found that compared with collectivistic societies, in individualistic societies, emotions feature more prominently in people's judgments because emotions provide direct feedback about the fit between personal needs and goals and the stimulus experienced (Schimmack and Diener 1997; Schimmack et al. 2002). In collectivistic societies, fulfillment of personal goals is also important, but relatively more weight is given to following cultural norms. In these cultures, there is relatively more emphasis on acquiring skills and gaining expertise than in individualistic societies (Hofstede 2001; Schimmack et al. 2002). Consequently, we expect that emotional experience receives more weight in the value derived from a Web site visit in individualistic societies than in collectivistic societies, whereas the utilitarian experience is given more emphasis in collectivistic societies than in individualistic societies.

In individualistic societies, the social fabric and group norms are much looser. These societies emphasize the private self, whereas the social self receives relatively more weight in collectivistic societies (Trafimow, Triandis, and Goto 1991). Individualistic societies emphasize that "everyone has a right to a private life," whereas in collectivistic cultures, "private life is invaded by institutions and organizations to which one belongs" (Hofstede 2001, p. 227). Personal security is also given more weight in individualistic cultures because the mechanism operant in collectivistic societies—namely, that security is provided by the organization or society—is relied on less in individualistic societies (Hofstede 2001, pp. 226–27, 267). In summary, it can be expected that personal privacy and security are given more weight in individualistic societies.

In their seminal article, Oyserman, Coon, and Kemmelmeier (2002) conclude that desire for uniqueness and independence are core elements of cultural individualism. In individualistic societies, a person's attitudes and behavior are regulated largely by individual preferences, whereas in collectivistic societies, attitudes and behavior are heavily influenced by society's preferences (Triandis 1989). In individualistic countries, children are socialized to think of themselves as "I." That is, an individual's identity is clearly distinct from other people's "I's" (Hofstede 2001, p. 227). People in individualistic societies value consumption experiences that are customized to their own unique needs more than people in collectivistic societies. In summary, we propose the following moderating effects involving national-cultural individualism:

H_5: The effect of utilitarian experience on perceived value is stronger when national-cultural individualism is lower.

H_6: The effect of emotional experience on perceived value is stronger when national-cultural individualism is higher.

H_7: The effect of perceived privacy/security protection on perceived value is stronger when national-cultural individualism is higher.

H_8: The effect of perceived customization on perceived value is stronger when national-cultural individualism is higher.

Method

Data Collection

Data collection was carried out in close collaboration with 16 U.S. and European CPG companies. They allowed us to post a questionnaire on their Web sites in one or more of the following countries: Belgium, France, Germany, Italy, the Netherlands, Spain, Sweden, and the United Kingdom. We obtained information on the individual-level constructs using questionnaires. We used secondary sources for the country characteristics. Furthermore, we conducted in-depth interviews with ten consumers to augment the quantitative results with qualitative insights into why the Web sites provided high or low perceived value. These consumers were not aware of the results of the large-scale quantitative study.

Questionnaires. We posted a survey on 30 Web sites and requested that online visitors fill out the survey. Electronic data collection was conducted in collaboration with the global market research agencies Taylor Nelson Sofres and GfK and the Internet market research agency MetrixLab. Online questionnaires were completed by 9546 consumers in 107 countries. We excluded 370 respondents because they had an excessive number of missing responses. In addition, for this study, we used only the countries from which at least 15 responses were returned. This rule resulted in the exclusion of an additional 290 completed surveys from further analysis. Our final data set consisted of 8886 consumers from 23 countries (Australia, Austria, Belgium, Canada, Croatia, Denmark, France, Germany, Ireland, Italy, Macedonia, the Netherlands, New Zealand, Poland, Romania, Russia, Serbia and Montenegro, Slovenia, Spain, Sweden, Switzerland, the United Kingdom, and the United States) reporting on 30 Web sites of 16 CPG companies. The list of companies and Web sites examined appears in Table 2.

Secondary data. We obtained country data from external sources (Hofstede 2001; ISSP 1998; Kaufmann, Kraay, and Zoido-Lobatón 1999; Suh et al. 1998).

Measures

We developed the questionnaire in English in close collaboration with an advisory committee of Internet experts (e.g., vice president of Internet marketing) of five of the participating CPG companies. We subsequently discussed the questionnaire in field interviews with 14 senior managers from ten of the participating CPG companies and pretested it on a Web site. After we made the necessary modifications, the questionnaire was finalized, back-translated into

TABLE 2
Firms, Brands, Product Categories, and Web Sites Examined

Firm	Web Site	Brand	Product Category
Barilla	www.barilla.com	Barilla	Pasta and pasta sauces
	www.barilla.it	Barilla	Pasta and pasta sauces
Beiersdorf	www.nivea.de	Nivea	Personal care
Cadbury Schweppes	www.schweppeseuro.com	Schweppes	Beverages
Colgate-Palmolive	www.colgate-palmolive.de	Multiple brands (e.g., Colgate, Ajax)	Personal care, household care, pet food
Danone	www.danoneconseils.com	Danone, Lu	Dairy products, biscuits, beverages
Heineken	www.heineken.com	Heineken	Beer
Heinz	www.heinz.co.uk	Heinz	Ketchup, beans, spaghetti
Henkel	www.DonnaD.it	Dixan, Pril, Vernel, Testanera	Household care, personal care
Kraft Foods	www.milka.de	Milka	Chocolate confectionery
	www.milkaski.de	Milka	Chocolate confectionery
	www.jacobs.de	Jacobs	Coffee
	www.estrella.se	Estrella	Salty snacks
Lindt & Sprüngli	www.lindt.com	Lindt	Chocolate confectionery
Mars	www.mypetstop.com/fr	Pedigree, Whiskas	Pet food
	www.mypetstop.com/ge	Pedigree, Whiskas	Pet food
	www.celebrations.co.uk	Celebrations	Confectionery
Procter & Gamble	www.olay.com	Olay	Personal care
	www.uk.pantene.com	Pantene Pro V	Hair care
Reckitt Benckiser	www.finish.co.uk	Finish	Dishwashing products
Sara Lee	www.de.nl	Douwe Egberts	Coffee
	www.senseocrema.nl	Senseocrema	Coffee pads
	www.geschenkenwinkel.nl	Douwe Egberts, Pickwick, Senseo	Coffee, tea, tableware, textiles, coffee makers
	www.zwitsal.nl	Zwitsal	Baby care
SCA Hygiene	www.bodyform.co.uk	Bodyform	Sanitary napkins
Unilever	www.knorr-gastro.ch	Knorr	Food products
	www.conimex.nl	Conimex	Foreign food products
	www.lifeandcooking.nl	Multiple brands (e.g., Becel, Bertolli)	Food products
	www.viadirekt.nu	Via Direkt	Laundry detergents
	www.axe.be	Axe	Deodorizer

Notes The content of these sites has changed since we carried out our consumer survey.

all languages involved (Dutch, French, German, Italian, Spanish, Swedish), and linked to the respective country Web sites. To satisfy Web site managers' concerns, the number of items per construct was limited. This is not uncommon in Internet research conducted in collaboration with companies (cf. Shankar, Smith, and Rangaswamy 2003). The Appendix provides detailed information on the measures.

Individual-level constructs. We measured perceived value with five items, which collectively capture the perceptions and behavioral intentions that have been associated with perceived value in the literature (Parasuraman, Zeithaml, and Malhotra 2005; Sirdeshmukh, Singh, and Sabol 2002; see also Holbrook 1999; Chen and Wells 1999). Seven facets of utilitarian experience are distinguished: (1) information provision, (2) easier decision making, (3) consumer education, (4) decision confidence, (5) convenience, (6) cross-channel integration, and (7) timeliness. These seven facets constitute a formative scale because there is no a priori reason to assume that they are correlated (Diamantopoulos and Winklhofer 2001). The higher a site scores on any of these facets in the mind of a visitor, the higher is his or her utilitarian experience. We measured both pleasure and arousal with two items taken from the work of Mehrabian and Russell (1974). We measured perceived privacy/security protection with two items, which we adapted from the work of Wolfinbarger and Gilly (2003) and Zeithaml, Parasuraman, and Malhotra (2002). We measured customization with three items based on the work of Srinivasan, Anderson, and Ponnavolu (2002) and Wolfinbarger and Gilly (2003).

Finally, we operationalized the congruity between the consumer and the site as a dummy variable (1 = site is adapted to the respondent's culture; 0 = otherwise). Adapted sites have three characteristics: (1) They are displayed in local country-specific language (e.g., American versus British English); (2) they offer local content (e.g., the site lists local activities versus the site is a mere translation); and (3) they exhibit some level of localization in the form of country-specific dates, zip codes, currency, and/or number formats (cf. Singh, Furrer, and Ostinelli 2004). Two coders independently coded whether each of the 30 sites was adapted to each of the 23 countries; there was a 96% interrater agreement, and disagreements were resolved by discussion.

Country-level constructs. Rule of law measures the respect of citizens and the state for the institutions that govern their interactions. It includes several indicators that measure the extent to which people have confidence in and abide by the rules of society. Together, these indicators measure the success of a society in developing an environment in which fair, predictable, impartial, and enforceable rules form the basis for economic and social interactions (Kaufmann, Kraay, and Zoido-Lobatón 1999). The rule-of-law measure reflects the statistical compilation of information obtained from surveys, nongovernmental organizations, commercial risk-rating agencies, and think tanks.

We based the national identity measure on the ISSP (1998) on national identity. Respondents in various countries were asked to rate their opinion on six topics pertaining to national identity, such as how close they felt to their respective countries, national pride, preference to their own country's films and programs, and willingness to move outside their country. The ISSP reports the country-averaged score for each item. We constructed a national identity scale by averaging the countries' scores on the six national identity items.[1]

We obtained the national-cultural individualism rating for each country by averaging the ratings of two leading experts in the field: Hofstede and Triandis, after converting Hofstede's ratings to a ten-point scale used by Triandis (e.g., Suh et al. 1998). Hofstede's scores have the strength of being based on a massive empirical data set, and Triandis's ratings include more recent information and include qualitative insights as well.

Covariates. We included two sociodemographic variables (gender, age), another site characteristic variable (accessibility), and 29 site dummy variables to control for unobserved site-specific effects (Greene 2000). We measured accessibility with two items pertaining to being able to access the site quickly and being able to maneuver through the site (Novak, Hoffman, and Yung 2000). These covariates are not the focus of our study, but controlling for their effects provides a stronger test of our hypotheses (Greene 2000).

Measure Validation

To allow for a stringent test of convergent and discriminant validity, we included all reflective individual-level constructs in a single confirmatory factor model. Although the chi-square was highly significant (not unexpected, given the large sample size), other indicators suggested a good fit (Byrne 1998): $\chi^2(67) = 2070.77$, $p < .001$; comparative fit index = .95; Tucker–Lewis index = .93; and root mean square error of approximation = .06. All the indicators loaded significantly ($p < .001$) and substantively (all factor loadings exceeded .50) on their hypothesized factors, providing evidence of convergent validity of the measures (Steenkamp and Van Trijp 1991). To test for discriminant validity, we ran a series of nested confirmatory factor analyses in which we constrained the correlation between each pair of reflective constructs to one. For all pairs, when we compared the constrained model with a free model, the difference was significant, which supports discriminant validity (Anderson and Gerbing 1988). Composite reliabilities were .65 or greater, exceeding the benchmarks that are suggested as acceptable (Bagozzi and Yi 1988).

We further tested the validity of the measures of the Web site characteristics perceived privacy/security protection and perceived customization by relating them to objective site characteristics. Drawing on the definition of privacy/security protection, along with consumer pretesting and discussions with the advisory committee of Internet experts of the participating CPG companies, we specified the following four objective privacy/security indicators: (1) whether the site tells how it will use the visitor's personal information, (2) whether the site explicitly indicates that it will not share personal data with third parties, (3) whether the site explicitly indicates that it will not send follow-up e-mails or other communications unless the visitor requests this, and (4) whether the visitor can delete the information the site has about him or her in its database.

In a similar vein, we developed five objective indicators for customization: (1) whether the visitor can personalize the content of the site, (2) whether the visitor can personalize the "look and feel" of the site, (3) whether the visitor can personalize fun activities and games on the site, (4) whether the visitor can consult an expert through the site, and (5) whether the visitor is addressed by name when visiting the site (after having registered). Two coders scored each Web site on each of these objective characteristics (1 = yes; 0 = no); there was a 97% interrater agreement, and disagreements were resolved by discussion. The correlation between the summated score on the objective privacy/security indicators and the mean perceived privacy/security score per site was .53 ($p < .001$), whereas this correlation was .50 for customization ($p < .001$). These correlations support the validity of the perceived Web site characteristics measures.

---

[1] For five countries, national identity ratings were unavailable from ISSP (1998). Instead, we used the national identity scores from a neighboring country. We conducted a robustness check by estimating the model after deleting those countries. The results remained substantively the same.

Model Specification

Our conceptual model of the antecedents of perceived value involves variables at two hierarchical levels of aggregation: the individual and the country level. The Level 1 (individual-level) and Level 2 (country-level) models are as follows:[2]

Level 1:

$$\begin{aligned}(1)\quad \text{PV}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{UTIL}_{ij} + \beta_{2j}\text{PLEAS}_{ij} + \beta_{3j}\text{AR}_{ij} \\ & + \beta_{4j}\text{AR}_{ij}^2 + \beta_{5j}\text{PRIVSEC}_{ij} \\ & + \beta_{6j}\text{CUSTOM}_{ij} + \beta_{7j}\text{CONGRUITY}_{ij} \\ & + \beta_{8j}\text{GENDER}_{ij} + \beta_{9j}\text{AGE}_{ij} + \beta_{10j}\text{ACCESS}_{ij} \\ & + \sum_{11}^{39} \beta_{qj}\text{SITE}_{ij} + r_{ij}.\end{aligned}$$

Level 2:

$$(2)\quad \beta_{0j} = \gamma_{00} + \gamma_{01}\text{LAW}_j + \gamma_{02}\text{NATID}_j + \gamma_{03}\text{INDIV}_j + u_{0j};$$

$$(3)\quad \beta_{1j} = \gamma_{10} + \gamma_{11}\text{INDIV}_j + u_{1j};$$

$$(4)\quad \beta_{2j} = \gamma_{20} + \gamma_{21}\text{INDIV}_j + u_{2j};$$

$$(5)\quad \beta_{3j} = \gamma_{30} + \gamma_{31}\text{INDIV}_j + u_{3j};$$

$$(6)\quad \beta_{5j} = \gamma_{50} + \gamma_{51}\text{LAW}_j + \gamma_{52}\text{INDIV}_j + u_{5j};$$

$$(7)\quad \beta_{6j} = \gamma_{60} + \gamma_{61}\text{INDIV}_j + u_{6j};$$

$$(8)\quad \beta_{7j} = \gamma_{70} + \gamma_{71}\text{NATID}_j + u_{7j};$$

$$(9)\quad \beta_{rj} = \gamma_{r0} + u_{rj}, \text{ for } r = 4, 8\text{–}10; \text{ and}$$

$$(10)\quad \beta_{sj} = \gamma_{s0}, \text{ for } s = 11, \ldots, 39;$$

where i and j represent individuals and countries, respectively; PV denotes perceived value; UTIL is utilitarian experience; PLEAS represents pleasure; AR is arousal; PRIVSEC measures perceived privacy/security protection; CUSTOM is perceived customization; CONGRUITY is cultural congruity; LAW is the rule of law; NATID is national identity; and INDIV is national-cultural individualism. The covariates GENDER and AGE are self-explanatory, ACCESS is the Web site's perceived accessibility, and SITE is the dummy variables that denote the specific Web site that was evaluated to control for unobserved site effects.

We assume the individual-level error term r_{ij} to be normally distributed with zero mean and variance σ^2. The random effects u_{qj} ($q = 0, \ldots, 10$) are multivariate normally distributed over countries, each with an expected value of zero, $\text{var}(u_{qj}) = \tau_{qq}$, and $\text{cov}(u_{qj}, u_{q'j}) = \tau_{qq'}$ ($q, q' = 0, \ldots, 10$). Finally, u_{qj} is the unique effect of country j on the intercept (β_{0j}) or slope ($\beta_{1j}, \ldots, \beta_{10j}$), while controlling for the country-level predictor variables (when applicable). A preliminary analysis indicated that our specification of random effects was consistent with the data. In all cases, the variance of the error term was significantly different from zero ($ps < .05$).

We do not hypothesize main effects for the country-level variables, because it is not obvious why the perceived value of a Web site is systematically influenced by, for example, the degree of national-cultural individualism of the country the visitor is from. However, we include these main effects as covariates for correct estimation of the interaction effects involving the country-level variables (Pedhazur 1982).

Model Estimation

In our model, respondents are nested within countries. Applying ordinary least squares regression to multilevel data is inappropriate because the resulting estimates will be biased and the estimated standard errors of the effects will be too small (Aitkin, Anderson, and Hinde 1981; Raudenbush 1988). Hierarchical linear modeling (HLM; Raudenbush and Bryk 2002) has been specifically developed to deal with multilevel data. It enables the simultaneous estimation of relationships of variables at two (or more) levels, using iterative maximum likelihood estimation. Hierarchical linear modeling borrows strength from all the data in each of the countries and makes it possible to estimate cross-level effects, thus enabling us to test H_3–H_8, specifying how country-level constructs affect relationships occurring at the individual level. We used HLM to test our research hypotheses. Following Raudenbush and Bryk's (2002) recommendations, we centered the Level 1 predictors within countries, and we grand-mean-centered the Level 2 predictors.

Results

Table 3 presents descriptive statistics and correlations between the variables. The results for the HLM model appear in Table 4.[3] Note that we report unstandardized coefficients. In HLM, standardized coefficients are problematic

---

[2] We estimated our model using SAS PROC MIXED (Singer 1998).

[3] We use one-tailed tests because we have directional hypotheses. A 5% significance level might be considered too lenient, given the large sample size, but we believe that this is not the case. It is well known that moderator effects are extremely difficult to detect in nonexperimental research (McClelland and Judd 1993). In simulation studies, McClelland and Judd (1993, p. 379) show that "91% of the simulated field studies made Type II errors by failing to reject a false null hypothesis." They argue (p. 387) that "enormous samples are required to have the statistical power of optimally designed experiments for detecting interactions."

TABLE 3
Descriptive Statistics

	Correlation Matrix												
	1	2	3	4	5	6	7	8	9	10	11	12	13
1 Perceived value	1.000												
2 Utilitarian experience	.563	1.000											
3 Pleasure	.446	.338	1.000										
4 Arousal	.460	.264	.659	1.000									
5 Privacy/security	.281	.309	.271	.250	1.000								
6 Customization	.540	.530	.471	432	.402	1.000							
7 Cultural congruity	011	.016	–.030	–.019	.015	–.012	1.000						
8 Gender	.155	.156	.067	.053	.096	.134	.003	1.000					
9 Age	–.009	–.032	–.042	.047	–.031	–.007	–.015	–.197	1.000				
10 Accessibility	.459	.472	.222	.196	.148	.311	014	.078	.001	1.000			
11 Rule of law	–.187	–.076	.011	–.190	–.111	–.108	.114	–.144	–.058	–.061	1.000		
12 National identity	–.048	.035	–.056	–.230	.056	.022	–.103	.072	–.141	–.058	–.033	1.000	
13 Individualism	–.137	–.015	.007	–.236	–.083	–.098	.045	–.032	–.099	– 059	.703	–.016	1.000
M	3.961	3.864	3.783	3.625	3.647	3.742	.951	.746	32.163	3 582	1.168	3.521	7.220
SD	.628	.438	.693	.761	.642	.639	.216	.435	12.096	.690	.858	.280	1.448

Notes: N = 8886 for evaluating pairwise correlations between individual-level variables or between individual- and country-level variables and for means and standard deviations of individual-level variables, N = 23 for evaluating pairwise correlations between country-level variables and for means and standard deviations of country-level variables.

because the variance is partitioned across different levels. We use t-values as a measure of relative effect size (Nijs et al. 2001).

The more positive the utilitarian experience ($\gamma_{10} = .421$, $p < .001$) and the more pleasurable the experience ($\gamma_{20} = .106$, $p < .001$), the higher is the perceived value. We also found an inverse U-shaped relationship between arousal and perceived value ($\gamma_{30} = .119$, $p < .001$; $\gamma_{40} = -.028$, $p < .01$). Thus, H_1 is supported. Contrary to our expectations (H_{2a} and H_{2c}), we do not find support for the notion that perceived privacy/security protection ($\gamma_{50} = .016$, $p = .11$) and cultural congruity ($\gamma_{70} = .017$, $p = .30$) enhance perceived value. However, as we expected (H_{2b}), customization has a significant effect on perceived value ($\gamma_{60} = .166$, $p < .001$).

The core of our conceptual model pertains to the cross-level interactions between individual-difference and country-level variables. The positive effect of perceived privacy/security protection on perceived value is stronger in countries characterized by a weak rule of law than in countries characterized by a strong rule of law ($\gamma_{51} = -.094$, $p < .01$), in support of H_3. The effect of perceived privacy/security protection is also larger in individualistic cultures than in collectivistic cultures ($\gamma_{52} = .040$, $p < .01$). This is consistent with H_7.

H_4 proposes that the effect of cultural congruity on perceived value is reinforced when national identity is higher. This hypothesis is supported ($\gamma_{71} = .198$, $p < .05$). H_5 proposes that the effect of utilitarian experience is smaller in individualistic cultures than in collectivistic cultures. This hypothesis is not supported ($\gamma_{11} = -.012$, $p = .32$). H_6, which proposes that the effect of emotional experience is larger in individualistic cultures than in collectivistic cultures, is supported for pleasure ($\gamma_{21} = .024$, $p < .05$) but not for arousal ($\gamma_{31} = -.001$, $p = .49$). Finally, as we expected, the effect of customization is larger when national-cultural individualism is higher ($\gamma_{61} = .035$, $p < .05$), in support of H_8.[4]

The main effects of gender and age were not significant. However, given the practical importance of these easily identifiable consumer characteristics, we reestimated our model including all two-way interactions with consumption experience and Web site characteristics. Three interactions were significant. We found an interaction between gender and $(\text{arousal})^2$ ($\gamma = -.033$, $p < .05$), indicating that males have a higher optimum level of arousal than females. This is consistent with the literature (Zuckerman 1994). Furthermore, we found that older people attach less importance to pleasure ($\gamma = -.003$, $p < .001$) and more importance to customization ($\gamma = .003$, $p < .001$).

Discussion

In this article, we examined whether the antecedents of the perceived value that consumers derive from visiting a brand

---

[4] As a robustness check, we reestimated the model for an alternative specification of perceived value, after eliminating Perceived Value Items 4 and 5, while adding Utilitarian Experience Item 3. The results remained substantively the same. The direction of all significant coefficients remained the same, but p values were often somewhat weaker. This is also because the alternative perceived value instrument had substantially lower reliability (.65 versus .75 for our operationalization). Nevertheless, nine of the ten hypothesized effects that were significant remained significant in the alternative model (the other coefficient attained a p value of .13).

TABLE 4
Results

Predictor	Hypothesized Effect	Unstandardized Coefficient	t-Value	p
Intercept (γ_{00})		4.033	80.2	<.001
Consumption Experience				
Utilitarian (γ_{10})	+	.421	14.8	<.001
Pleasure (γ_{20})	+	.106	8.3	<.001
Arousal (γ_{30})	+	.119	4.2	<.001
(Arousal)^2 (γ_{40})	–	–.028	–2.6	<.01
Web Site Characteristics				
Privacy/security (γ_{50})	+	.016	1.3	n.s.
Customization (γ_{60})	+	.166	7.4	<.001
Cultural congruity (γ_{70})	+	.017	.5	n.s.
Moderating Effect of Country Characteristics				
Rule of law × privacy/security (γ_{51})	–	–.094	–2.7	<.01
National identity × cultural congruity (γ_{71})	+	.198	2.0	<.05
Individualism × utilitarian experience (γ_{11})	–	–.012	–.5	n.s.
Individualism × pleasure (γ_{21})	+	.024	1.8	<.05
Individualism × arousal (γ_{31})	+	–.001	–.1	n.s.
Individualism × privacy/security (γ_{52})	+	.040	2.6	<.01
Individualism × customization (γ_{61})	+	.035	1.8	<.05
Covariates				
Gender (0 = male; 1 = female) (γ_{80})		.025	1.1	n.s.
Age (γ_{90})		–.001	–.4	n.s.
Accessibility (γ_{100})		.177	9.2	<.001
Rule of law (γ_{01})		–.076	–1.9	n.s.
National identity (γ_{02})		–.062	–.7	n.s.
Individualism (γ_{03})		.016	.8	n.s.
29 site dummies (γ_{110})–(γ_{390})				

Notes: Except for the covariates, all p values are for one-sided tests; n.s = not significant (p > .05).

manufacturer's Web site are systematically moderated by the country's institutional context. We developed specific hypotheses and tested them across 30 CPG Web sites in eight European countries, using large samples of consumers from three continents. We find that perceived privacy/security is more important in countries characterized by a weak rule of law. Perceived privacy/security and customization have a greater effect in individualistic countries than in collectivistic countries. Cultural congruity weighs more heavily in countries that are high on national identity. We find evidence that the emotional experience is more important in shaping perceived value judgments in individualistic countries, but this effect holds for pleasure only. Notably, existing research on the difference in the impact of emotions in people's judgments between individualistic and collectivistic societies largely concerns pleasure (e.g., Hofstede 2001; Schimmack et al. 2002; Triandis, McCusker, and Hui 1990). It appears that, in general, it is not so much emotions but rather the pursuit of pleasure that is more important in judgment formation in individualistic countries. Finally, there is directional evidence that utilitarian experience is more important in collectivistic societies, but this effect does not reach statistical significance.

In summary, our findings provide broad support for our conceptual model and the relevance of including country-level institutional constructs for understanding the perceived value that consumers derive from visiting a brand manufacturer's Web site. Our results show that the relationships between perceived Web site value and its drivers are systematically and predictably moderated by the institutional context in which consumers live. The findings of this study are important from a theoretical perspective because they offer insights into why certain variables do not significantly affect perceived value in some countries; namely, the national environment may not be conducive to such relationships. It is also important from an applied perspective because more and more companies use Web sites to create consumer value.

Managerial Implications

At the micro level, the most important driver of perceived value is the utilitarian experience associated with the site. A company can increase utilitarian experience by providing clear, truthful, and new information about the company's products in a way that eases decision making and increases decision confidence. Notably, some Web sites of these large CPG companies were not able to achieve this, at least not in the minds of their visitors. For example, in the follow-up in-depth interviews, consumers indicated two Web sites that received low scores on utilitarian experience (the consumers in the qualitative interviews were not informed about the results of the quantitative study), both of which were operated by major European CPG companies: "This site is just an advertisement instead of providing useful content," and "this site packs so many bells and whistles that the useful information gets lost."

On average, pleasure is the second most important driver of perceived value. Additional time and money spent on creating sites that generate stronger emotions of pleasure are particularly beneficial for sites that are primarily visited by people from individualistic countries. Visual material appears to be a major component of pleasurable experiences. A Web site operated by Sara Lee was able to create a pleasurable experience by using soft colors and cute animations, such as a small flying butterfly. In the follow-up qualitative interviews, consumers mentioned that "this site is really pleasant; it's like taking a walk in a fairy-tale wood." Visitors of Danone's (called Dannon in the United States) site explained that they considered the site visit a pleasurable experience mainly because of the nice colors and the high-quality pictures: "This is a pleasurable site to visit because of the high-quality pictures that are being used. I consider that to be very important"; and "this site looks pretty sunny because of the light colors that are being used; it makes one feel sunny too."

The third most important driver is customization. Especially people living in more individualistic countries gave more weight to customization in their perceived value judgments. Customization can be achieved by allowing the visitor to adapt the content, or the look and feel, of the site to his or her own preferences (e.g., by changing the color scheme at Henkel's site), by allowing the visitor to consult an expert (e.g., through a live consult at Procter & Gamble's Olay), or by personally addressing the visitor. After registration, visitors of Mars's Mypetstop can create a personalized page for each pet they have. They can use this personalized page to keep track of medical records and vet appointments and to set up an online scrapbook to post photos of their pet so other pet lovers and friends can view them. In these personalized pages, the visitor's pet is always referred to by its name rather than by the impersonal "your pet." In-depth interviews with consumers revealed that they particularly like the customization offered by the site: "This is a very personal site; it feels as if this site is directly talking to me and my cat"; and "my cat is not just an animal on this site; it is given a personality, and it is addressed by its name." In addition, visitors can use the FoodFinder to find out what pet food product would be advisable for their pet after they have typed in the pet's type, breed, age, special needs, and whether it is overweight.

The inverted U-shaped effect of arousal generalized across countries. Our qualitative interviews indicated that a major way to generate arousal is to engage the Web site visitor actively. Two examples are Barilla and Unilever's Life & Cooking. The qualitative interviews indicated the following: "The strength of this site is the way its recipe section is set up; it is definitely fun to play around with" (Barilla); and "I feel excited because the site invites you to actively use it by allowing you to collect your own favorite recipes into a personal cookery book, create menus for special dinners, e-mail these to yours friends/guests, and print out shopping lists" (Life & Cooking).

The effect of perceived privacy/security protection was completely moderated by the country context. Especially Web sites oriented to countries in which individualism is high and the rule of law is weak should emphasize privacy/security. Our measurement validation check suggests that this can be achieved by informing visitors about how the site intends to use the visitors' personal information, by explicitly indicating that personal data will not be shared with third parties and that follow-up e-mails will not be sent, and by making it clear to visitors that they can always delete the data the site has about them in its marketing database. Our qualitative interviews further revealed that in addition to legally formulated privacy/security statements, a well-known brand name can also act as a privacy/security signal. Illustrative of this are the following comments, referring to Kraft Foods's brand Milka: "I trust an old brand like Milka, so I would not hesitate when asked to hand over my credit card number"; and "they're a well-known brand, and they clearly indicate a telephone number where you can call them when you have questions or something goes wrong, so why should I worry?" Consumers' reliance on a reputable brand for privacy/security protection is consistent with research showing that relying on a well-known brand name is a frequently employed risk reduction strategy (Roselius 1971; see also Sultan et al. 2002) and with economic theory suggesting that consumers can plausibly believe in the protection offered by a high-reputation brand name (Choi 1998; Shapiro 1983; Wernerfelt 1988).

The effect of cultural congruity was also completely moderated by the country context. We find that Web sites oriented to countries in which national identity is high should be adapted to the local context by using local language, symbols, and so forth. The global marketing literature posits that completely standardized products often have limited market appeal, whereas complete adaptation to the local context creates cost inefficiencies and may result in products that are not better than completely standardized products (e.g., because the wheel is reinvented every time and research-and-development expenditures are spread too thinly; Kotabe and Helsen 2004). The Internet is a powerful instrument to both globalize and localize. Companies can use global portals to direct Web site visitors to nationally tailored sites. This is consistent with received wisdom in global marketing to develop a common "core," while allowing adaptation to local circumstances. A case in point is Beiersdorf's site for Nivea. Nivea's portal gives its users access to more than 40 regional and country sites; they share the same look and feel, including Nivea's signature blue colors, and typically feature Innovations, Magazine, Advice, and Fun as the main buttons. However, the sites are locally adapted with respect to language, the model used (e.g., an Asian woman for sites in Japan, a Western woman for Western sites), the innovations featured on the country portal, and so forth. The Internet offers unrivaled potential to adapt Web sites and company communications locally around a globally standardized platform, thus reaping the efficiencies of a standardized product and the responsiveness of local adaptation (Yip 2003).

Limitations and Further Research

Our research has several limitations that offer avenues for further research. First, although our study would have been difficult to carry out without the close collaboration of major CPG companies, this also imposed constraints on the design of the study. These companies were interested in

MEASUREMENT APPENDIX

Individual-Level Constructs

Perceived Value (α = .75)

1. Taken everything together, I consider this Web site useful.
2. This Web site fell short of my expectations. (R)
3. I do not intend to bookmark this Web site (add to favorites). (R)
4. I intend to visit this Web site again.
5. Visiting this Web site was a poor experience. (R)

Utilitarian Experience (Formative Scale)

This site ...

1. Provides clear and truthful information about products.
2. Makes it more difficult to decide what to choose. (R)
3. Allows me to learn useful things I did not know before.
4. Lowers my confidence to make the right choices. (R)
5. Makes it more difficult to fulfill my shopping needs. (R)
6. Indicates stores where I can get additional information or buy the product.
7. Provides information that is up-to-date.

Pleasure (α = .70)

When surfing this site ...

1. I feel pleased.
2. I feel happy.

Arousal (α = .70)

When surfing this site ...

1. I feel stimulated.
2. I feel excited.

Perceived Privacy/Security Protection (α = .66)

1. This site has a rigorous privacy policy.
2. This site has good protection against fraud.

Perceived Customization (α = .65)

This site ...

1. Has interactive features that fit my needs.
2. Helps me to evaluate offerings that fit my needs.
3. Creates the feeling of receiving personalized attention.

Cultural Congruity

= 1 if the site is adapted to the respondent's culture.
= 0 if otherwise.

Accessibility (Formative Scale; Covariate)

1. Downloading information from this site is slow. (R)
2. It is difficult to find the information I need on this site. (R)

Country-Level Constructs

National Cultural Individualism

We employed a combination of country scores from the work of Hofstede (2001) and Suh and colleagues (1998) and converted them into a ten-point scale.

National Identity (α = .89)

1. If you could improve your work or living conditions, how willing or unwilling would you be to move outside your country? (1 = "very willing"; 5 = "very unwilling")
2. I would rather be a citizen of my country than of any other country in the world. (1 = "agree strongly"; 5 = "disagree strongly") (R)
3. The world would be a better place if people from other countries were more like the people in my country. (1 = "agree strongly"; 5 = "disagree strongly") (R)
4. Generally my country is a better country than most other countries. (1 = "agree strongly"; 5 = "disagree strongly") (R)
5. My country should limit the import of foreign products in order to protect its national economy. (1 = "agree strongly"; 5 = "disagree strongly") (R)
6. My country's television should give preference to my country's films and programs. (1 = "agree strongly"; 5 = "disagree strongly") (R)

Rule of Law

We employed the country ratings provided in the work of Kaufmann, Kraay, and Zoido-Lobatón (1999).

Notes: R = reverse-scored item. With the exception of cultural congruity, all individual-level items were scored using a five-point scale with the following categories: "strongly disagree," "disagree," "neither agree nor disagree," "agree," and "strongly agree."

using Web sites as a communication tool for creating perceived value for consumers. Further research could investigate other communication objectives (e.g., creating brand value) and examine Web sites that offer an (additional) online sales channel.

Second, consistent with the perspective of the participating CPG companies, our study takes the perspective of the consumer by focusing on the consumption experience associated with the Web site visit and perceived Web site characteristics. Although the follow-up in-depth interviews with consumers and the measurement validation highlighted some of the ways that managers can create the desired consumer experiences and perceptions, further research could systematically examine what causes Web sites to deliver these experiences and perceptions. A promising possibility is to content-analyze Web sites more extensively and relate objective site characteristics to the variables in our conceptual model.

Third, we assessed the measure of cultural congruity on a dichotomous scale. Further research could measure this on a continuous scale. Finally, the consumers in our sample visited the Web sites on their own accord. This approach seemed appropriate because our study is about explaining variance in perceived value. Further research might contrast perceptions of visitors versus nonvisitors.

REFERENCES

Aitkin, Murray, Dorothy Anderson, and John Hinde (1981), "Statistical Modeling of Data on Teaching Styles," *Journal of the Royal Statistical Society, Series A*, 144 (4), 419–61.

Allik, Juri and Anu Realo (2004), "Individualism-Collectivism and Social Capital," *Journal of Cross-Cultural Psychology*, 35 (January), 29–49.

Anderson, James C. and David W Gerbing (1988), "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin*, 103 (3), 411–23.

Arndt, Johan (1979), "Towards a Concept of Domesticated Markets," *Journal of Marketing*, 43 (Fall), 69–75.

Bagozzi, Richard P. and Youjae Yi (1988), "On the Evaluation of Structural Equation Models," *Journal of the Academy of Marketing Science*, 16 (Spring), 74–94.

Balabanis, George, Adamantios Diamantopoulos, Rene D. Mueller, and T.C. Melewar (2001), "The Impact of Nationalism, Patriotism, and Internationalism on Consumer Ethnocentric Tendencies," *Journal of International Business Studies*, 32 (1), 157–75.

Baughn, C. Christopher and Attila Yaprak (1996), "Economic Nationalism: Conceptual and Empirical Development," *Political Psychology*, 17 (4), 759–78.

Berlyne, Daniel E. (1971), *Aesthetics and Psychobiology*. Englewood Cliffs, NJ: Prentice Hall.

Billig, Michael (1995), *Banal Nationalism*. London: Sage Publications.

Byrne, Barbara M. (1998), *Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS: Basic Concepts, Applications, and Programming*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chen, Qimei, Sandra J. Clifford, and William D. Wells (2002), "Attitude Toward the Site II: New Information," *Journal of Advertising Research*, 42 (March–April), 33–45.

——— and William D. Wells (1999), "Attitude Toward the Site," *Journal of Advertising Research*, 39 (September–October), 27–37

Choi, Jay P. (1998), "Brand Extension as Informational Leverage," *Review of Economic Studies*, 65 (4), 655–69.

Curry, Sheree R. (1999), "Kraft Site Shows CPG Firms How It's Done," *Marketing News*, (October 11), 20.

Day, George S. (1990), *Market Driven Strategies: Processes for Creating Value*. New York: The Free Press.

Diamantopoulos, Adamantios and Heidi M. Winklhofer (2001), "Index Construction with Formative Indicators: An Alternative to Scale Development," *Journal of Marketing Research*, 38 (May), 269–77.

Druckman, Daniel (1994), "Nationalism, Patriotism, and Group Loyalty: A Social-Psychological Perspective," *Mershon International Studies Review*, 38 (1), 43–68.

Eliashberg, Jehoshua and Mohanbir S. Sawhney (1994), "Modeling Goes to Hollywood: Predicting Individual Differences in Movie Enjoyment," *Management Science*, 40 (9), 1151–73.

Feldman, Jack M. and John G. Lynch (1988), "Self-Generated Validity and Other Effects of Measurement on Belief, Attitude, Intention, and Behavior," *Journal of Applied Psychology*, 73 (3), 421–35.

Gale, Bradley T. (1994), *Managing Customer Value*. New York: The Free Press

Giddens, Anthony (1981), *A Contemporary Critique of Historical Materialism*. London: McMillan.

——— (1984), *The Constitution of Society*. Berkeley: University of California Press.

Greene, William H. (2000), *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall.

Grewal, Rajdeep and Ravi Dharwadkar (2002), "The Role of the Institutional Environment in Marketing Channels," *Journal of Marketing*, 66 (July), 82–97

Hoffman, Andrew W. (2001), *From Heresy to Dogma: An Institutional History of Corporate Environmentalism*. Palo Alto, CA: Stanford University Press.

Hoffman, Donna L., Thomas P. Novak, and M.A. Peralta (1999), "Building Consumer Trust Online," *Communications of the ACM*, 42 (4), 80–85.

Hofstede, Geert (2001), *Culture's Consequences*. Thousand Oaks, CA: Sage Publications.

Holbrook, Morris B. (1994), "The Nature of Customer Value," in *Service Quality*, Roland T. Rust and Richard L. Oliver, eds. Thousand Oaks, CA. Sage Publications, 21–71.

——— (1999), *Consumer Value: A Framework for Analysis and Research*. New York· Routledge.

——— and Kim P. Corfman (1985), "Quality and Value in the Consumption Experience: Phaedrus Rides Again," in *Perceived Quality*, Jacob Jacoby and Jerry C Olson, eds. Lexington, MA: Lexington Books, 31–57.

——— and Elizabeth C Hirschman (1982), "The Experiential Aspects of Consumption· Consumer Fantasies, Feelings, and Fun," *Journal of Consumer Research*, 9 (September), 132–40.

ISSP (1998), *International Social Survey Program· National Identity*. Ann Arbor, MI: Inter-University Consortium for Political and Social Research.

Kagitcibasi, Cigdem (1997), "Individualism and Collectivism," in *Handbook of Cross-Cultural Psychology: Social Behavior and Applications*, Vol. 3, John W. Berry, Marshall H. Segall, and Cigdem Kagitcibasi, eds Boston: Allyn & Bacon, 1–49

Kaufmann, Daniel, Aart Kraay, and Pablo Zoido-Lobatón (1999), "Aggregating Governance Indicators," Policy Research Paper No. 2195, The World Bank.

Koslow, Scott, Prem N. Shamdasani, and Ellen E. Touchstone (1994), "Exploring Language Effects in Ethnic Advertising: A Sociolinguistic Perspective," *Journal of Consumer Research*, 20 (December), 575–85.

Kotabe, Masaaki and Kristiaan Helsen (2004), *Global Marketing Management*. Hoboken, NJ: John Wiley & Sons.

Kowalski, Robin M. and Randall Wolfe (1994), "Collective Identity Orientation, Patriotism, and Reactions to National Outcomes," *Personality and Social Psychology Bulletin*, 20 (5), 533–40.

Lastovicka, John L. (1982), "On the Validation of Lifestyle Traits: A Review and Illustration," *Journal of Marketing Research*, 19 (February), 126–38.

Licht, Amir N., Chanan Goldschmidt, and Shalom H. Schwartz (2003), "Culture Rules: The Foundations of the Rule of Law and Other Norms of Governance," William Davidson Institute Working Paper Series 2003-605, Stephen M. Ross Business School, University of Michigan.

Luna, David, Laura A. Peracchio, and Maria D. de Juan (2002), "Cross-Cultural and Cognitive Aspects of Web Site Navigation," *Journal of the Academy of Marketing Science*, 30 (4), 397–410.

Lwin, May O. and Jerome D. Williams (2004), "A Model Integrating the Multidimensional Developmental Theory of Privacy and Theory of Planned Behavior to Examine Fabrication of Information Online," *Marketing Letters*, 14 (4), 257–72.

Mackie, Diane M. and Eliot R. Smith (1998), "Intergroup Relations: Insights from a Theoretically Integrative Approach," *Psychological Review*, 105 (3), 499–529.

McClelland, Gary H. and Charles M. Judd (1993), "Statistical Difficulties of Detecting Interactions and Moderator Effects," *Psychological Bulletin*, 114 (2), 376–90.

Mehrabian, Albert and James Russell (1974), *An Approach to Environmental Psychology*. Cambridge, MA: MIT Press.

Naumann, Earl (1995), *Creating Customer Value*. Cincinnati: Thompson.

New Media Age (2003), "15% of Top FMCG Brands Lack a Web Site," (April 3), 14.

Nijs, Vincent R., Marnik G. Dekimpe, Jan-Benedict E.M. Steenkamp, and Dominique H. Hanssens (2001), "The Category-Demand Effects of Price Promotions," *Marketing Science*, 20 (Winter), 1–22.

North, Douglass C. (1990), *Institutions, Institutional Change and Economic Performance*. Cambridge, UK: Cambridge University Press.

Novak, Thomas P., Donna L. Hoffman, and Yiu-Fai Yung (2000), "Measuring the Customer Experience in Online Environments: A Structural Modeling Approach," *Marketing Science*, 19 (1), 22–42.

Olney, Thomas J., Morris B. Holbrook, and Rajeev Batra (1991), "Consumer Responses to Advertising: The Effects of Ad Content, Emotions, and Attitude Toward the Ad on Viewing Time," *Journal of Consumer Research*, 17 (December), 440–53.

Oxley, Joanne E. and Bernard Yeung (2001), "E-Commerce Readiness: Institutional Environment and International Competitiveness," *Journal of International Business Studies*, 32 (4), 705–723.

Oyserman, Daphna, Heather M. Coon, and Markus Kemmelmeier (2002), "Rethinking Individualism and Collectivism: Evaluation of Theoretical Assumptions and Meta-Analyses," *Psychological Bulletin*, 128 (1), 3–72.

Parasuraman, A., Valarie A. Zeithaml, and Arvind Malhotra (2005), "E-S-QUAL: A Multiple-Item Scale for Assessing Electronic Service Quality," *Journal of Service Research*, 7 (3), 213–33.

Pedhazur, Elazar J. (1982), *Multiple Regression in Behavioral Research*. New York: Holt, Rinehart and Winston.

Raudenbush, Stephen W. (1988), "Educational Applications of Hierarchical Linear Models. A Review," *Journal of Educational Statistics*, 13 (Summer), 85–116.

——— and Anthony S. Bryk (2002), *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, CA. Sage Publication.

Reisenzein, Rainer (1994), "Pleasure-Arousal Theory and the Intensity of Emotions," *Journal of Personality and Social Psychology*, 76 (3), 525–39.

Roselius, Ted (1971), "Consumer Rankings of Risk Reduction Methods," *Journal of Marketing*, 35 (January), 56–61.

Schimmack, Ulrich and Ed Diener (1997), "Affect Intensity: Separating Intensity and Frequency in Repeatedly Measured Affect," *Journal of Personality and Social Psychology*, 73 (6), 1313–29.

———, Phanikirian Radhakrishnan, Shigehiro Oishi, Vivian Dzokoto, and Stephan Ahadi (2002), "Culture, Personality, and Subjective Well-Being: Integrating Process Models of Life Satisfaction," *Journal of Personality and Social Psychology*, 82 (4), 582–93.

Schwartz, Shalom H. (1994), "Beyond Individualism/Collectivism: New Cultural Dimensions of Value," in *Individualism and Collectivism: Theory, Method, and Applications*, U. Kim, H.C. Triandis, C. Kagitcibasi, S.C Choi, and G. Yoon, eds. Thousand Oaks, CA: Sage Publications, 85–119.

Scott, W. Richard (2001), *Institutions and Organizations*. Thousands Oaks, CA. Sage Publications.

Shankar, Venkatesh, Amy K. Smith, and Arvind Rangaswamy (2003), "Customer Satisfaction and Loyalty in Online and Offline Environments," *International Journal of Research in Marketing*, 20 (2), 153–75.

Shapiro, Carl (1983), "Premiums for High Quality Products as Returns to Reputation," *Quarterly Journal of Economics*, 98 (4), 659–80.

Shimp, Terence A. and Subhash Sharma (1987), "Consumer Ethnocentrism: Construction and Validation of the CETSCALE," *Journal of Marketing Research*, 24 (August), 280–89.

Singer, Judith D. (1998), "Using SAS PROC MIXED to Estimate Multilevel Models, Hierarchical Models, and Individual Growth Models," *Journal of Educational and Behavioral Statistics*, 24 (4), 323–55.

Singh, Nitish, Oliver Furrer, and Massimiliano Ostinelli (2004), "To Localize or to Standardize on the Web: Empirical Evidence from Italy, India, Netherlands, Spain, and Switzerland," *Multinational Business Review*, 12 (1), 69–87.

Sirdeshmukh, Deepak, Jagdip Singh, and Barry Sabol (2002), "Consumer Trust, Value, and Loyalty in Relational Exchanges," *Journal of Marketing*, 66 (January), 15–37.

Srinivasan, Srini S., Rolph Anderson, and Kishore Ponnavolu (2002), "Customer Loyalty in E-Commerce: An Exploration of Its Antecedents and Consequences," *Journal of Retailing*, 78 (1), 41–50.

Steenkamp, Jan-Benedict E.M., Rajeev Batra, and Dana L. Alden (2003), "How Perceived Brand Globalness Creates Brand Value," *Journal of International Business Studies*, 34 (1), 53–65

——— and Hans C.M. van Trijp (1991), "The Use of LISREL in Validating Marketing Constructs," *International Journal of Research in Marketing*, 8 (4), 283–99.

Suchman, Mark C. (1995), "Managing Legitimacy: Strategic and Institutional Approaches," *Academy of Management Review*, 20 (July), 571–610.

Suh, Eunkook, Ed Diener, Shigehiro Oishi, and Harry C. Triandis (1998), "The Shifting Basis of Life Satisfaction Judgments Across Cultures: Emotions Versus Norms," *Journal of Personality and Psychology*, 74 (2), 482–93.

Sultan, Fareena, Glen L. Urban, Venkatesh Shankar, and Iakov Y. Bart (2002), "Determinants and Role of Trust in E-Business: A

Large Scale Empirical Study," working paper, College of Business Administration, Northeastern University.

Tajfel, Henri (1981), *Human Groups and Social Categories*. Cambridge, UK: University Press.

Tedeschi, Bob (2003), "E-Commerce Report: Consumer Products Companies Use Web Sites to Strengthen Ties with Customers," *The New York Times*, (August 25), 6.

Trafimow, David, Harry C. Triandis, and Sharon G. Goto (1991), "Some Tests of the Distinction Between the Private Self and the Collective Self," *Journal of Personality and Social Psychology*, 60 (5), 649–55.

Triandis, Harry C. (1989), "The Self and Social Behavior in Differing Cultural Contexts," *Psychological Review*, 96 (3), 506–520.

——— (1995), *Individualism and Collectivism*. Boulder, CO: Westview Press.

———, Christopher McCusker, and C. Harry Hui (1990), "Multimethod Probes of Individualism and Collectivism," *Journal of Personality and Social Psychology*, 59 (5), 1006–1020.

Verlegh, Peeter W.J. (2001), *Country-of Origin Effects on Consumer Product Evaluations*, doctoral dissertation, Marketing and Consumer Behaviour Group, Wageningen University, the Netherlands.

Wells, William D. and David Prensky (1996), *Consumer Behavior*. New York: John Wiley & Sons.

Wernerfelt, Birger (1988), "Umbrella Branding as a Signal of New Product Quality: An Example of Signaling by Posting a Bond," *Rand Journal of Economics*, 19 (3), 458–66.

Wolfinbarger, Mary and Mary C. Gilly (2003), ".eTailQ: Dimensionalizing, Measuring and Predicting eTail Quality," *Journal of Retailing*, 79 (3), 183–98.

Woodruff, Robert B. (1997), "Customer Value: The Next Source of Competitive Advantage," *Journal of the Academy of Marketing Science*, 25 (2), 139–53.

Yip, George S. (2003), *Total Global Strategy*. Upper Saddle River, NJ: Prentice Hall.

Zeithaml, Valarie A. (1988), "Consumer Perceptions of Price, Quality, and Value: A Conceptual Model and Synthesis of Research," *Journal of Marketing*, 52 (July), 2-22.

———, A. Parasuraman, and Arvind Malhotra (2002), "Service Quality Delivery Through Web Sites: A Critical Review of Extant Knowledge," *Journal of the Academy of Marketing Science*, 30 (4), 362–375.

Zuckerman, Marvin D. (1994), *Behavioral Expressions and Biosocial Bases of Sensation Seeking*. New York: Cambridge University Press.

Journal of Public Policy & Marketing
Fall 2007

Call for Papers

The American Marketing Association's New Definition of Marketing: Perspectives on Its Implications for Scholarship and the Role and Responsibility of Marketing in Society

Special Section Editor: Gregory T. Gundlach, University of North Florida

Since 1935, as part of its role, the American Marketing Association (AMA) has offered to both academics and practitioners its version of the definition of marketing. These definitions are widely referenced by academics and used by marketing practitioners. In 2004, AMA announced a new definition of marketing: *Marketing is an organizational function and a set of processes for creating, communicating and delivering value to customers and for managing customer relationships in ways that benefit the organization and its stakeholders.* The new definition was unveiled at the 2004 AMA Summer Educators' Conference and replaces its prior definition. According to AMA, this definition incorporates the contributions of many marketers from around the world, both academics and practitioners. Given the importance of a definition, including for the field of marketing, a special "section" of *Journal of Public Policy & Marketing* (*JPP&M*) will examine the implications of AMA's new definition of marketing for (1) scholarship in marketing and society and (2) the role and responsibility of marketing practice in society. Topics of interest for the special section include the following:

The implications of AMA's new definition for academic marketing:

- For the academic disciplines of marketing and society, including public policy, consumer interest, social marketing, ethics and macromarketing, and so forth.
- For scholarship within these disciplines, including knowledge development, dissemination, application, and integration.
- For education within these disciplines.
- For the future development and impact of these disciplines.
- For the relationship and impact of these disciplines on other fields.
- Domestic and international perspectives on the preceding points.

The implications of AMA's new definition for marketing practice:

- For the role of marketing practice in society.
- For the responsibility of marketing practice in society.
- For the future development and impact of the practice of marketing in society.
- For the relationship and impact of marketing practice on other professions.
- Domestic and international perspectives on the preceding points.

Additional topics of interest for the special section include:

- Assessment of the nature and meaning of definitions in fields of study and professions including marketing and marketing and society.
- Historical overviews of definitions of marketing and their relationship to marketing and society.
- Identification of dimensions and attributes to be included in a definition of marketing in relation to the field of marketing and society.
- Prediction of the impact of AMA's new definition for the status and reputation of the field of marketing and marketing and society.

JPP&M encourages the submission of "essays" on these and related topics that address the implications of AMA's new definition of marketing for scholarship in marketing and society and the role and responsibility of marketing practice in society. Papers will be reviewed in accordance with *JPP&M* guidelines. Submitted essays should not exceed 10–12 double-spaced pages, including references, tables, and figures (12-point font). Submissions (Word format) should be sent by e-mail to Gregory T. Gundlach, special section editor, at ggundlac@unf.edu by December 1, 2006.

Take A Walk into American Households

Meet the Smiths, and the Jones, and most every other consumer, too. The Yellow Pages brings your brand directly into American homes, reaching hundreds of millions of consumers when they're ready to buy. And 88% of YP users actually do – spending 25% more than consumers overall.[1] See for yourself. YP's payout is truly trackable and totally measurable, with a national sales ROI of 27:1.[2] Put this unstoppable business generator to work for you. Reach consumers in print, online, and wireless. Then, get ready. Almost 40% of YP users become new customers.[3] With this kind of reach, the message is clear: YP delivers. For every brand. And every national marketing plan.

Potent. Pervasive. Profitable. And you thought you knew YP. Call 1-866-88BUYYP (882-8997). Or visit www.buyyellow.com. It pays. We'll prove it.

YELLOW PAGES
ASSOCIATION
PRINT•INTERNET•WIRELESS

[1] KN/SRI Industry Usage Study, 2005 CRM/Simmons (based on 90 categories), 2005 [2] CRM Associates, Metered Ad Study of National Display Ads in Selected Directories 2005 [3] KN/SRI Industry Usage Study, 2005 ©2006, Yellow Pages Association All Rights Reserved

Journal of Marketing

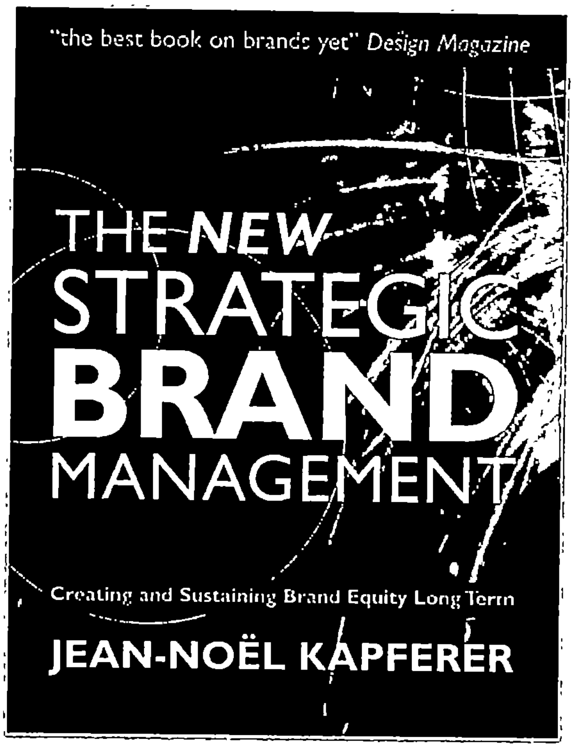
October 2006
Volume 70, Number 4

A Quarterly Publication of the
American Marketing Association

Corporate Social Responsibility, Customer Satisfaction, and Market Value	1	Xueming Luo & C.B. Bhattacharya
The Influence of Avatars on Online Consumer Shopping Behavior	19	Martin Holzwarth, Chris Janiszewski, & Marcus M. Neumann
Creating a Market Orientation: A Longitudinal, Multifirm, Grounded Analysis of Cultural Transformation	37	Gary F. Gebhardt, Gregory S. Carpenter, & John F. Sherry Jr.
It Just Feels Good: Customers' Affective Response to Touch and Its Influence on Persuasion	56	Joann Peck & Jennifer Wiggins
Marketing's Credibility: A Longitudinal Investigation of Marketing Communication Productivity and Shareholder Value	70	Xueming Luo & Naveen Donthu
Upgrades and New Purchases	92	Erica Mina Okada
Influence Tactics for Effective Adaptive Selling	103	Richard G. McFarland, Goutam N. Challagalla, & Tasadduq A. Shervani
How Biased Household Inventory Estimates Distort Shopping and Storage Decisions	118	Pierre Chandon & Brian Wansink
Factors Influencing the Effectiveness of Relationship Marketing: A Meta-Analysis	136	Robert W. Palmatier, Rajiv P. Dant, Dhruv Grewal, & Kenneth R. Evans
Achieving Marketing Objectives Through Social Sponsorships	154	Carolyn J. Simmons & Karen L. Becker-Olsen
The Unhealthy = Tasty Intuition and Its Effects on Taste Inferences, Enjoyment, and Choice of Food Products	170	Rajagopal Raghunathan, Rebecca Walker Naylor, & Wayne D. Hoyer
Broadening the Scope of Reference Price Advertising Research: A Field Study of Consumer Shopping Involvement	185	Daniel J. Howard & Roger A. Kerin
Subject and Author Index to Volume 70	205	

AMERICAN MARKETING ASSOCIATION

One of the cornerstones of brand management in MBA programs today


"the best book on brands yet" Design Magazine
THE NEW STRATEGIC BRAND MANAGEMENT
Creating and Sustaining Brand Equity Long Term
JEAN-NOËL KAPFERER

PAPER, 520 PP, 0 7494 4283 2, $59.95

- ☐ New models and benchmarks for building brands and businesses
- ☐ New approach to brand identity and positioning, brand architecture and multi-brand portfolios
- ☐ Addresses key issues, including:
  - – *Building brands in saturated markets*
  - – *Managing retail brands*
  - – *Energizing brand value chains*
  - – *Brand extensions*
  - – *Brand innovation*
  - – *Launching brands*
  - – *Globalizing existing brands*
  - – *Revitalizing aging brands*

JN Kapferer (HEC, Ph.D Northwestern University) is considered one of world's top branding experts. He has been a consultant for European, Asian and American companies and teaches advanced executive development seminars worldwide. In 2004, he received the AMA's S. Taumer Cavusgil Award.

KOGAN PAGE

For more details, visit **www.koganpage.com** *or contact us for an inspection copy:*

525 South 4th Street, Suite 241 • Philadelphia, PA 19147
Telephone: 215-928-9112 or 215-525-3251
Fax 215-928-9113 • E-mail: info@koganpage.com

Journal of Marketing

Vol. 70, No. 4

AMERICAN MARKETING ASSOCIATION

California University
11/6/07
CAMPUS

EDITOR
Roland T. Rust
Robert H. Smith School of Business
3451 Van Munching Hall
University of Maryland
College Park, MD 20742-1815

ASSISTANT TO THE EDITOR
Marti Tiedeman
E-mail: jom@ama.org
Tel: (312) 542-9024

Millicent Locke
University of Maryland

WEB EDITOR
Siva K. Balasubramanian
Southern Illinois University

MARKETING LITERATURE REVIEW
Myron Leonard
Western Carolina University

BOOK REVIEWS
Ramarao Desiraju
University of Central Florida

VICE PRESIDENT—PUBLICATIONS
Russell S. Winer
New York University

CHIEF EXECUTIVE OFFICER
Dennis Dunlap

GROUP PUBLISHER
Jack Hollfelder

EDITORIAL STAFF
Phone: (312) 542-9000
Fax: (312) 542-9001

MANAGING EDITOR
Francesca Van Gorp Cooley

TECHNICAL EDITOR
Andy Seagram

GRAPHICS EDITOR
Mary Loye

Advertising Staff
Fax: (312) 922-3763

Richard Ballschmiede, Advertising
(312) 542-9076

Sally Schmitz, Production
(312) 542-9038

JM Web site:
Submissions: www.marketingjournals.org/JM
Main page: www.marketingpower.com/JM

The *Journal of Marketing* (ISSN 0022-2429) is published quarterly (January, April, July, and October). *Annual Subscription Rates*: Subscriptions are $95 for individuals in the United States and Canada and $120 in other countries. Library, school, and business subscriptions are $235 in the United States and Canada and $265 in other countries. Subscriptions for AMA members are $55. Prices are for one year. Single copy rate: $50 individual, $60 corporate in the United States and Canada; add $5 in other countries Prepayment is required in U.S. dollars or the equivalent.

Published by the Publications Group of the American Marketing Association, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266. Periodicals postage paid at Chicago, IL and additional mailing offices. POSTMASTER: Send address changes to the *Journal of Marketing*, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266, USA.

•MANUSCRIPTS: Electronic submissions must be made via the *JM* Web site at http://www.marketingjournals.org. Paper submissions are not accepted Editorial policies, review board members, Tables of Contents for *JM*, and so on, are available on the *JM* Web site at http://www.marketingpower.com/jm.

•ADVERTISERS AND ADVERTISING AGENCIES assume liability for all content (including text, representations, and illustrations) of advertisements published, and also assume responsibility for any claims arising therefrom made against the publisher. The right is reserved to reject any advertisement.

•SUBSCRIPTIONS*:* Members: Send to American Marketing Association, 311 S. Wacker Dr, Suite 5800, Chicago, IL 60606-2266 (1-800-AMA-1150 or 312/542-9000), or e-mail to info@ama.org. Nonmembers· Send to American Marketing Association, P.O Box 11806, Birmingham, AL 35202, 1-800-633-4931, 205-995-1567, Fax· 205-995-1588, e-mail· amasubs@ebsco.com. Canadian residents add 6% GST (GST registration # 12747852).

•PERMISSIONS AND REPRINT INFORMATION: Without written permission from the AMA, any copying or reprinting (except by authors reprinting their own work) is prohibited. Requests for permission to reprint—such as copying for general distribution, advertising, or promotional purpose; for creating new collective works; or for resale—should be submitted in writing to permissions@ama.org. Bulk reprints are available by contacting Tamara Groft, Sheridan Reprints, (800) 352-2210. Permission is granted by AMA for those registered with the Copyright Clearance Center (CCC), 222 Rosewood Dr, Danvers, MA 01923, to photocopy any article herein for a flat fee Contact CCC with ISSN 0022-2429, volume, and first and last page numbers of each article copied.

The views expressed in *JM* are not necessarily those of the editorial staff or the publisher. **All articles in this journal, unless otherwise noted, have undergone a blind review process.**

Copyright © 2006, American Marketing Association. All rights reserved.

Editorial Review Board

ROHINI AHLUWALIA
University of Minnesota

KUSUM L AILAWADI
Dartmouth College

JOSEPH ALBA
University of Florida

TIM AMBLER
London Business School

EUGENE W. ANDERSON
University of Michigan

TOR WALLIN ANDREASSEN
Norwegian School of Management

ASIM ANSARI
Columbia University

ERIC J. ARNOULD
University of Arizona

KWAKU ATUAHENE-GIMA
CEIBS

BARRY L BAYUS
University of North Carolina

WILLIAM BEARDEN
University of South Carolina

PAUL D BERGER
Boston University

PAUL N BLOOM
University of North Carolina

LISA E. BOLTON
University of Pennsylvania

RUTH N. BOLTON
Arizona State University

ROBERT BORDLEY
General Motors Corporation

DOUGLAS BOWMAN
Emory University

RODERICK J. BRODIE
University of Auckland

STEPHEN P BROWN
University of Houston

JOSEPH CANNON
Colorado State University

DIPANKAR CHAKRAVARTI
University of Colorado

PIERRE CHANDON
INSEAD

YUXIN CHEN
New York University

PRADEEP K CHINTAGUNTA
University of Chicago

KEVIN CLANCY
Copernicus Marketing Consulting

PETER J DANAHER
University of Auckland

JOHN A DEIGHTON
Harvard University

MARNIK G DEKIMPE
Catholic University Leuven

WAYNE S DESARBO
Pennsylvania State University

RAVI DHAR
Yale University

NAVEEN DONTHU
Georgia State University

TULIN ERDEM
University of California, Berkeley

PETER S FADER
University of Pennsylvania

GARY L FRAZIER
University of Southern California

INGE GEYSKENS
Tilburg University

PETER N GOLDER
New York University

RAJDEEP GREWAL
Pennsylvania State University

ABBIE GRIFFIN
University of Illinois, Urbana-Champaign

SUNIL GUPTA
Columbia University

JOHN R HAUSER
Massachusetts Institute of Technology

JAN B. HEIDE
University of Wisconsin, Madison

CHRISTIAN HOMBURG
University of Mannheim

WAYNE D. HOYER
University of Texas at Austin

G. TOMAS M. HULT
Michigan State University

J JEFFREY INMAN
University of Pittsburgh

GANESH IYER
University of California, Berkeley

CHRIS JANISZEWSKI
University of Florida

SANDY D. JAP
Emory University

BERNARD J. JAWORSKI
Monitor Group

MICHAEL D. JOHNSON
Cornell University

PK KANNAN
University of Maryland

TIMOTHY L KEININGHAM
Ipsos Loyalty

AMNA KIRMANI
University of Maryland

AJAY K KOHLI
Emory University

ROBERT V. KOZINETS
York University

V KUMAR
University of Connecticut

DONALD R LEHMANN
Columbia University

KATHERINE N LEMON
Boston College

MARY FRANCES LUCE
Duke University

ROBERT F LUSCH
University of Arizona

JOHN G LYNCH JR.
Duke University

DEBORAH J. MACINNIS
University of Southern California

VIJAY MAHAJAN
University of Texas at Austin

CARL F MELA
Duke University

VIKAS MITTAL
University of Pittsburgh

WENDY MOE
University of Maryland

CHRISTINE MOORMAN
Duke University

JOHN MOWEN
Oklahoma State University

EITAN MULLER
New York University and Tel-Aviv University

SCOTT A NESLIN
Dartmouth College

RICHARD NETEMEYER
University of Virginia

STEPHEN M. NOWLIS
Arizona State University

JOSEPH C. NUNES
University of Southern California

A PARASURAMAN
University of Miami

KOEN PAUWELS
Dartmouth College

WILLIAM D PERREAULT
University of North Carolina

ROBERT PETERSON
University of Texas at Austin

F.G M. PIETERS
Tilburg University

JAIDEEP PRABHU
Imperial College

ARVIND RANGASWAMY
Pennsylvania State University

VITHALA RAO
Cornell University

BRIAN T RATCHFORD
University of Texas at Dallas

AMERICUS REED
University of Pennsylvania

WERNER REINARTZ
INSEAD

ARIC RINDFLEISCH
University of Wisconsin-Madison

FRED SELNES
Norwegian School of Management

STEVEN M SHUGAN
University of Florida

ITAMAR SIMONSON
Stanford University

REBECCA SLOTEGRAAF
Indiana University

DILIP SOMAN
University of Toronto

JOYDEEP SRIVASTAVA
University of Maryland

JAN-BENEDICT E M STEENKAMP
University of North Carolina

DAVID W STEWART
University of Southern California

STEFAN STREMERSCH
Erasmus University Rotterdam

BAOHUNG SUN
Carnegie Mellon University

GERARD J. TELLIS
University of Southern California

JACQUELYN THOMAS
Northwestern University

CRAIG THOMPSON
University of Wisconsin-Madison

STIJN VAN OSSELAER
Erasmus University Rotterdam

P. RAJAN VARADARAJAN
Texas A&M University

STEPHEN L VARGO
University of Hawaii, Manoa

PETER VERHOEF
University of Groningen

KENNETH WATHNE
University of Wisconsin-Madison

MICHEL WEDEL
University of Maryland

BARTON A. WEITZ
University of Florida

RUSSELL S. WINER
New York University

ANTHONY ZAHORIK
Burke Institute

VALARIE A ZEITHAML
University of North Carolina, Chapel Hill

Z. JOHN ZHANG
University of Pennsylvania

Ad Hoc Reviewers

The people listed here worked jointly with the Editorial Review Board by devoting their time and specialized knowledge to reviewing a substantial number of manuscripts submitted to JM between July 1, 2005, and June 30, 2006. The Editor greatly appreciates their help in the review process.

RAVI S ACHROL
George Washington University

MARJORIE ADAMS
Morgan State University

JAGDISH AGRAWAL
California State University
Hayward

LERZAN AKSOY
Koç University

IAN ALAM
State University of New York at
Geneseo

GERALD ALBAUM
University of New Mexico

DANA L ALDEN
University of Hawaii, Honolulu

MARK I. ALPERT
University of Texas at Austin

ERIC T ANDERSON
Northwestern University

ERIN ANDERSON
INSEAD

PER E. ANDERSSON
Stockholm School of Economics

ALAN R. ANDREASEN
Georgetown University

J. CRAIG ANDREWS
Marquette University

JONLEE ANDREWS
Indiana University

KERSI D. ANTIA
University of Wisconsin–Madison

J. SCOTT ARMSTRONG
University of Pennsylvania

NEERAJ ARORA
University of Wisconsin–Madison

EMIN BABAKUS
University of Memphis

BARRY J. BABIN
University of Southern Mississippi

RICHARD P. BAGOZZI
University of Michigan

WILLIAM E. BAKER
San Diego State University

SRIDHAR BALASUBRAMANIAN
University of North Carolina at
Chapel Hill

RAVI BAPNA
University of Connecticut

MICHAEL J BARONE
Iowa State University

BRUCE BARRY
Vanderbilt University

DEBRA Z BASIL
University of Lethbridge

SUMAN BASUROY
Florida Atlantic University

RAJEEV BATRA
University of Michigan

RONALD J. BAUERLY
Western Illinois University

HANS BAUMGARTNER
Pennsylvania State University

SHARON BEATTY
University of Alabama

RUSSELL W BELK
University of Utah

DANIEL C. BELLO
Georgia State University

NEELI BENDAPUDI
Ohio State University

LANCE A. BETTENCOURT
Indiana University

SUNDAR G BHARADWAJ
Singapore Management University

C. B. BHATTACHARYA
Boston University

BARBARA BICKART
Rutgers University

GABRIEL J BIEHAL
University of Maryland

MARY JO BITNER
Arizona State University

EYAL BIYALOGORSKY
University of California, Davis

PETER H. BLOCH
University of Missouri–Columbia

JAMES S. BOLES
Georgia State University

PAULA FITZGERALD BONE
West Virginia University

PAUL A. BOTTOMLEY
Cardiff University

BILL BOULDING
Duke University

ERIC T BRADLOW
University of Pennsylvania

MICHAEL K BRADY
Florida State University

THOMAS G BRASHEAR
University of Massachusetts
Amherst

SUSAN M. BRONIARCZYK
University of Texas at Austin

BART J. BRONNENBERG
University of California, Los
Angeles

JAMES R BROWN
West Virginia University

STEPHEN W BROWN
Arizona State University

TOM J. BROWN
Oklahoma State University

MATTHIAS BUCERIUS
Merck KGaA

PETER J. BURKE
University of California, Riverside

RAYMOND R. BURKE
Indiana University

ALVIN C. BURNS
Louisiana State University

SCOT BURTON
University of Arkansas

PAUL S. BUSCH
Texas A&M University

ROGER J CALANTONE
Michigan State University

DENNIS CAMPBELL
Harvard University

MARGARET C CAMPBELL
University of Colorado

NOEL CAPON
Columbia University

LES CARLSON
Clemson University

ZIV CARMON
INSEAD

GREGORY S. CARPENTER
Northwestern University

S TAMER CAVUSGIL
Michigan State University

PAUL K CHANEY
Vanderbilt University

PATRALI CHATTERJEE
Rutgers University

SUBIMAL CHATTERJEE
State University of New York at
Binghamton

ARJUN CHAUDHURI
Fairfield University

JEAN-CHARLES W CHEBAT
HEC, Montreal

AMAR CHEEMA
Washington University in St. Louis

YUBO CHEN
University of Arizona

ALEXANDER CHERNEV
Northwestern University

TERRY L CHILDERS
University of Kentucky

PETER C.Y. CHIU
University of Cincinnati

MARKUS CHRISTEN
INSEAD

ALKA V CITRIN
Georgia Institute of Technology

DEL COATES
San Jose State University

CHARLES COLBY
Rockbridge Associates Inc

CATHERINE COLE
University of Iowa

MARK R. COLGATE
University of Victoria

T BETTINA CORNWELL
University of Queensland

JOSEPH A. COTE
Washington State University

ROBIN HIGIE COULTER
University of Connecticut

NICOLE E COVIELLO
University of Auckland

WILLIAM L CRON
Texas Christian University

J. JOSEPH CRONIN
Florida State University

PEGGY H CUNNINGHAM
Queen's University

SANDOR F. CZELLAR
HEC, Paris

PETER A DACIN
Queen's University

DARREN W. DAHL
University of British Columbia

ROBERT DAHLSTROM
University of Kentucky

FRANCESCA DALL'OLMO RILEY
Kingston University Business
School

ASWATH DAMODARAN
New York University

CHRYSANTHOS DELLAROCAS
University of Maryland

KO DE RUYTER
Maastricht University

KALPESH KAUSHI DESAI
State University of New York at
Buffalo

PREYAS DESAI
Duke University

ROHIT DESHPANDÉ
Harvard University

RAMARAO DESIRAJU
University of Central Florida

TIMOTHY M. DEVINNEY
University of New South Wales

RAVI DHARWADKAR
Syracuse University

ANIRUDH DHEBAR
Babson College

RUBY R. DHOLAKIA
University of Rhode Island

UTPAL M. DHOLAKIA
Rice University

ADAMANTIOS
DIAMANTOPOULOS
University of Vienna

C. ANTHONY DI BENEDETTO
Temple University

PETER R. DICKSON
Florida International University

KRISTIN DIEHL
University of Southern California

WILLIAM R DILLON
Southern Methodist University

ANDREA L DIXON
University of Cincinnati

D. TODD DONAVAN
Colorado State University

SUSAN P DOUGLAS
New York University

GRAHAME DOWLING
University of New South Wales

JAMES DREW
Verizon Laboratories

TOM DUNCAN
University of Denver

SHANTANU DUTTA
University of Southern California

F. ROBERT DWYER
University of Cincinnati

SEVGIN EROGLU
Georgia State University

JENNIFER EDSON ESCALAS
Vanderbilt University

HOOMAN ESTELAMI
Fordham University

KENNETH R. EVANS
University of Missouri–Columbia

RONALD FABER
University of Minnesota

LIAM FAHEY
Babson College

PAUL W FARRIS
University of Virginia

O. C. FERRELL
University of Wyoming

MARC FISCHER
University of Kiel

ROBERT J FISHER
University of Western Ontario

VALERIE S. FOLKES
University of Southern California

CLAES FORNELL
University of Michigan

NATASHA Z. FOUTZ
University of Maryland

PHILIP HANS B.F FRANSES
Erasmus University Rotterdam

JUDY K FRELS
University of Maryland

SHANKAR GANESAN
University of Arizona

JAISHANKAR GANESH
University of Central Florida

ELLEN GARBARINO
Case Western Reserve University

JAMES W GENTRY
University of Nebraska–Lincoln

MARY C GILLY
University of California, Irvine

FUSUN F GONUL
Füsun Consulting

GERALD J. GORN
Hong Kong University of Science and Technology

ROBERT GRANT
Georgetown University

KENT GRAYSON
Northwestern University

ERIC A. GREENLEAF
New York University

DWAYNE D. GREMLER
Bowling Green State University

DHRUV GREWAL
Babson College

CHRISTIAN GRONROOS
Hanken Swedish School of Economics, Finland

THOMAS S. GRUCA
University of Iowa

THOMAS GRUEN
University of Colorado at Colorado Springs

EVERT N. GUMMESSON
Stockholm University

GREGORY T GUNDLACH
University of North Florida

ZEYNEP GURHAN-CANLI
University of Michigan

ANDERS B. GUSTAFSSON
Karlstad University

REBECCA W HAMILTON
University of Maryland

JIN K HAN
Singapore Management University

DOMINIQUE M. HANSSENS
University of California, Los Angeles

BRUCE HARDIE
London Business School

CHARLES P. HARRIS
University of Otago

LLOYD C. HARRIS
Cardiff University

JAMES HARTLEY
Keele University

ERNAN HARUVY
University of Texas at Dallas

GERALD HAUBL
University of Alberta

DIANA L HAYTKO
Missouri State University

TIMOTHY B HEATH
Miami University

DAVID H. HENARD
North Carolina State University

THORSTEN HENNIG-THURAU
Bauhaus-University of Weimar

ANDREAS HERRMANN
University of St Gallen

BAS HILLEBRAND
Radboud University Nijmegen

ELIZABETH C HIRSCHMAN
Rutgers University

MICHAEL A HITT
Texas A&M University

STEPHEN J. HOCH
University of Pennsylvania

STEVE E. HOEFFLER
University of North Carolina at Chapel Hill

DONNA L. HOFFMAN
Vanderbilt University

JAN H HOFMEYR
Synovate

MORRIS B. HOLBROOK
Columbia University

MARIA HOLMLUND-RYTKONEN
Swedish School of Economics

DOUG HOLT
University of Oxford

MARK B HOUSTON
University of Missouri–Columbia

JOEL HUBER
Duke University

JOHN S HULLAND
University of Pittsburgh

ROBERT F HURLEY
Fordham University

MICHAEL R. HYMAN
New Mexico State University

SUBIN IM
San Francisco State University

ROBERT JACOBSON
University of Washington

WILLIAM G. JACOBY
Michigan State University

EUGENE D. JAFFE
Copenhagen Business School

SANJAY JAIN
Texas A&M University

WOLFGANG JANK
University of Maryland

SATISH JAYACHANDRAN
University of South Carolina

KAMEL JEDIDI
Columbia University

GITA VENKATARAMANI JOHAR
Columbia University

MARK C. JOHLKE
Bradley University

DEBORAH ROEDDER JOHN
University of Minnesota

GEORGE JOHN
University of Minnesota

KISSAN JOSEPH
University of Kansas

ANNAMMA JOY
Concordia University

VRINDA KADIYALI
Cornell University

MANOHAR U. KALWANI
Purdue University

WAGNER A KAMAKURA
Duke University

DIMITRI KAPELIANIS
Emory University

SUSAN M KEAVENEY
University of Colorado at Denver

SCOTT W KELLEY
University of Kentucky

KAREN NORMAN KENNEDY
University of Alabama at Birmingham

ROGER A. KERIN
Southern Methodist University

NAMWOON KIM
Hong Kong Polytechnic University

AHMET H. KIRCA
Michigan State University

RAN KIVETZ
Columbia University

JILL G. KLEIN
University of Texas at Austin

RICHARD H. KOLBE
Kent State University

PRAVEEN K. KOPALLE
Dartmouth College

MICHEL KOSTECKI
University of Neuchatel

BOB KRAPFEL
University of Maryland

ROBERT KRIDER
Simon Fraser University

BALAJI KRISHNAN
University of Memphis

VENKAT R. KRISHNAN
XLRI Jamshedpur

DEAN M KRUGMAN
University of Georgia

SABINE KUESTER
University of Mannheim

NIRMALYA KUMAR
London Business School

KYRIAKOS KYRIAKOPOULOS
ALBA Graduate Business School

BARBARA A. LAFFERTY
University of South Florida

VICKI R LANE
University of Colorado at Denver

RIKARD LARSSON
Lund University

GILLES LAURENT
HEC, Paris

ANGELA Y LEE
Northwestern University

MICHAEL LEVY
Babson College

DONALD R. LICHTENSTEIN
University of Colorado at Boulder

GARY L. LILIEN
Pennsylvania State University

YONG LIU
Syracuse University

CHRISTOPHER LOVELOCK
Yale University

BRYAN A LUKAS
University of Melbourne

XUEMING LUO
University of Texas at Arlington

RICHARD J LUTZ
University of Florida

SCOTT B. MACKENZIE
Indiana University

DAVID P. MACKINNON
Arizona State University

DURAIRAJ MAHESWARAN
New York University

NARESH K. MALHOTRA
Georgia Institute of Technology

EDWARD C. MALTHOUSE
Northwestern University

ELLIOT MALTZ
Willamette University

PUNEET MANCHANDA
University of Chicago

NAOMI MANDEL
Arizona State University

DETELINA MARINOVA
Case Western Reserve University

INGRID M MARTIN
California State University, Long Beach

HOWARD PAUL MARVEL
Ohio State University

CHARLOTTE H MASON
University of North Carolina

CHARLA K MATHWICK
Portland State University

KEN MATSUNO
Babson College

ANNA S MATTILA
Pennsylvania State University

JAMES G. MAXHAM
University of Virginia

JAMES H. MCALEXANDER
Oregon State University

LEIGH MCALISTER
University of Texas at Austin

EDWARD F. MCQUARRIE
Santa Clara University

JOHN T MENTZER
University of Tennessee

PAUL R. MESSINGER
University of Alberta

MATTHEW L. MEUTER
California State University, Chico

ROBERT J. MEYER
University of Pennsylvania

DAVID GLEN MICK
University of Virginia

PETER MILLS
University of Oregon

GEORGE R. MILNE
University of Massachusetts

DEBI PRASAD MISHRA
State University of New York at Binghamton

ANDREW A MITCHELL
University of Toronto

SUNIL K MITHAS
University of Maryland

DEBANJAN MITRA
University of Florida

BANWARI MITTAL
Northern Kentucky University

NATALIE MIZIK
Columbia University

JAKKI J. MOHR
University of Montana

MICHAEL P. MOKWA
Arizona State University

ELIZABETH S. MOORE
University of Notre Dame

KAJ PN MOREL
Delft University

NEIL A. MORGAN
Indiana University

MAUREEN A MORRIN
Rutgers University

ALBERT M. MUNIZ
DePaul University

JANET Y MURRAY
University of Missouri–St. Louis

KYLE B. MURRAY
University of Western Ontario

A V. MUTHUKRISHNAN
Hong Kong University of Science and Technology

NIKLAS MYHR
Chapman University

SRIDHAR NARAYANAN
Stanford University

DAS NARAYANDAS
Harvard University

NADA I. NASR
Bentley College

DHANANJAY NAYAKANKUPPAM
University of Iowa

BRUCE I. NEWMAN
DePaul University

EDWIN J. NIJSSEN
Radboud University Nijmegen

CHARLES H NOBLE
University of Mississippi

THOMAS P NOVAK
Vanderbilt University

NATHAN NOVEMSKY
Yale University

ARNE NYGAARD
Norwegian School of Management

HERBJORN NYSVEEN
Norwegian School of Economics and Business Administration

CARL OBERMILLER
Seattle University

GINA C. O'CONNOR
Rensselaer Polytechnic Institute

THOMAS C O'GUINN
University of Illinois at Urbana–Champaign

ROGELIO OLIVA
Texas A&M University

ERIC M OLSON
University of Colorado at Colorado Springs

AMY L OSTROM
Arizona State University

MICHAEL OSTROVSKY
Stanford University

PADDY PADMANABHAN
INSEAD

XING PAN
Indiana University

C. WHAN PARK
University of Southern California

YOUNG-HOON PARK
Cornell University

MARK E. PARRY
University of Virginia

VANESSA M. PATRICK
University of Georgia

PAUL PATTERSON
University of New South Wales

CORNELIA PECHMANN
University of California, Irvine

LISA PENALOZA
University of Colorado at Boulder

MICHEL TUAN PHAM
Columbia University

NIGEL F PIERCY
University of Warwick

JOHN W. PRACEJUS
University of Alberta

LINDA L PRICE
University of Arizona

S. DOUGLAS PUGH
University of North Carolina at Charlotte

DANIEL S. PUTLER
University of British Columbia

WILLIAM P PUTSIS
University of North Carolina at Chapel Hill

WILLIAM J QUALLS
University of Illinois at Urbana–Champaign

SONJA RADAS
Institute of Economics, Zagreb

PRIYA RAGHUBIR
University of California, Berkeley

V. KASTURI RANGAN
Harvard University

AKSHAY R RAO
University of Minnesota

AMBAR G. RAO
Washington University in St. Louis

RAM C. RAO
University of Texas at Dallas

S. ABRAHAM RAVID
Rutgers University

LOPO L REGO
University of Iowa

PETER H. REINGEN
Arizona State University

NANCY M. RIDGWAY
University of Richmond

NORA J. RIFON
Michigan State University

WILLIAM T ROBINSON
Purdue University

MICHELLE L ROEHM
Wake Forest University

JOSE ANTONIO ROSA
University of Illinois at Chicago

JOHN R. ROSSITER
University of Wollongong

HERBERT JACK ROTFELD
Auburn University

ROBERT I. ROUNDTREE
Howard University

GARY J. RUSSELL
University of Iowa

RICHARD M. RYAN
University of Rochester

HENRIK H SATTLER
University of Hamburg

LISA K. SCHEER
University of Missouri–Columbia

BODO B. SCHLEGELMILCH
Wirtschaftsuniversitat Wien

ANN E. SCHLOSSER
University of Washington

JOHN SCHOUTEN
University of Portland

SEETHU SEETHARAMAN
Rice University

KATHLEEN SEIDERS
Boston College

SANKAR SEN
Baruch College, The City University of New York

RAJESH SETHI
Clarkson University

RAJ SETHURAMAN
Southern Methodist University

VENKATESH SHANKAR
Texas A&M University

ARUN SHARMA
University of Miami

DANIEL A. SHEININ
University of Rhode Island

JOHN F. SHERRY
University of Notre Dame

TERENCE A. SHIMP
University of South Carolina

JORGE M. SILVA-RISSO
University of California, Riverside

JAGDIP SINGH
Case Western Reserve University

RAJIV K SINHA
Arizona State University

DEEPAK SIRDESHMUKH
North Carolina State University

STANLEY F. SLATER
Colorado State University

AMY K SMITH
George Washington University

CRAIG SMITH
London Business School

ALINA SORESCU
Texas A&M University

GILVAN C SOUZA
University of Maryland

MARTIN O SPANN
University of Passau

JOHN R SPARKS
University of Dayton

RICHARD SPEED
University of Melbourne

ROSANN L SPIRO
Indiana University

DAVID E. SPROTT
Washington State University

RAJI SRINIVASAN
University of Texas at Austin

RICHARD STAELIN
Duke University

THOMAS F. STAFFORD
University of Memphis

BARBARA B. STERN
Rutgers University

BRIAN STERNTHAL
Northwestern University

PATRICIA A. STOUT
University of Texas at Austin

VANITHA SWAMINATHAN
University of Pittsburgh

TERESA A. SWARTZ
California Polytechnic State University

NADER T. TAVASSOLI
London Business School

STEPHEN S. TAX
University of Victoria

STEVEN A. TAYLOR
Illinois State University

DAVID K. TSE
University of Hong Kong

MICHAEL TSIROS
University of Miami

RAJEEV K. TYAGI
University of California, Irvine

MARK D. UNCLES
University of New South Wales

H RAO UNNAVA
Ohio State University

JOEL E. URBANY
University of Notre Dame

CHRISTOPHE VAN DEN BULTE
University of Pennsylvania

MARC VANHUELE
HEC School of Management

KOERT VAN ITTERSUM
Georgia Institute of Technology

SAJEEV VARKI
Boston College

RAJKUMAR VENKATESAN
University of Connecticut

ALLADI VENKATESH
University of California, Irvine

R VENKATESH
University of Pittsburgh

FRANZISKA VOLCKNER
University of Hamburg

DOUGLAS W. VORHIES
University of Mississippi

GLENN B. VOSS
North Carolina State University

JANET WAGNER
University of Maryland

FLORIAN WANGENHEIM
University of Dortmund

BRIAN WANSINK
Cornell University

LUC WATHIEU
Harvard University

CHARLES B. WEINBERG
University of British Columbia

KLAUS WERTENBROCH
INSEAD

J CHRIS WHITE
Michigan State University

BEREND WIERENGA
Erasmus University Rotterdam

IAN F. WILKINSON
University of New South Wales

PATRICIA WILLIAMS
University of Pennsylvania

MARY FINLEY WOLFINBARGER
California State University, Long Beach

JOHN P WORKMAN
Creighton University

STEFAN WUYTS
Tilburg University

JINHONG XIE
University of Florida

SHA YANG
New York University

CATHERINE W. M. YEUNG
National University of Singapore

EDEN YIN
University of Cambridge

GAL ZAUBERMAN
University of North Carolina at Chapel Hill

FLORIAN ZETTELMEYER
University of California, Berkeley

Outstanding Reviewers

The *Journal of Marketing* was again the top-rated marketing journal for 2005, based on the SSCI impact factors, and was ranked second among *all* business journals. Maintaining this level of quality is only possible with fast, constructive, insightful reviews by the Editorial Review Board and ad hoc reviewers. The performance of the journal's reviewers for the year 2005/2006 has been truly outstanding. The average decision turnaround for new submissions for the most recent quarter was only 52 days, despite submissions to *JM* being at record levels (371 new submissions in 05/06). Among all the outstanding Editorial Review Board members, five people stood out as having more reviews than average, completing them faster than average, and having the highest level of quality. They are the winners of the ***Best Reviewer Award*** for 2005/2006:

Tim Ambler (London Business School)
John Hauser (MIT)
Donald Lehmann (Columbia)
Werner Reinartz (INSEAD)
Barton Weitz (Florida)

Congratulations to these outstanding reviewers!

Guest Editors

The *Journal of Marketing* thanks the following people for serving as guest editors during 2005/2006:

Ruth Bolton
Thomas Kinnear
Robert Lusch
David Stewart
P. Rajan Varadarajan

The journal is also pleased to announce 20 new appointments to the Editorial Review Board. The new appointees are:

Joseph Alba (Florida)
Kwaku Atuahene-Gima (CEIBS)
William Bearden (South Carolina)
Stephen P. Brown (Houston)
Joseph Cannon (Colorado State)
Yuxin Chen (NYU)
Inge Geyskens (Tilburg)
Wendy Moe (Maryland)
John Mowen (Oklahoma State)
Eitan Muller (NYU & Tel Aviv)
Richard Netemeyer (Virginia)
Koen Pauwels (Dartmouth)
Robert Peterson (Texas)
Vithala Rao (Cornell)
Aric Rindfleisch (Wisconsin)
Fred Selnes (Norwegian School of Management)
Rebecca Slotegraaf (Indiana)
Jacqueline Thomas (Northwestern)
Craig Thompson (Wisconsin)
Kenneth Wathne (Wisconsin)

Xueming Luo & C.B. Bhattacharya

Corporate Social Responsibility, Customer Satisfaction, and Market Value

Although prior research has addressed the influence of corporate social responsibility (CSR) on perceived customer responses, it is not clear whether CSR affects market value of the firm. This study develops and tests a conceptual framework, which predicts that (1) customer satisfaction partially mediates the relationship between CSR and firm market value (i.e., Tobin's q and stock return), (2) corporate abilities (innovativeness capability and product quality) moderate the financial returns to CSR, and (3) these moderated relationships are mediated by customer satisfaction. Based on a large-scale secondary data set, the results show support for this framework. Notably, the authors find that in firms with low innovativeness capability, CSR actually reduces customer satisfaction levels and, through the lowered satisfaction, harms market value. The uncovered mediated and asymmetrically moderated results offer important implications for marketing theory and practice.

In today's competitive market environment, corporate social responsibility (CSR) represents a high-profile notion that has strategic importance to many companies. As many as 90% of the *Fortune* 500 companies now have explicit CSR initiatives (Kotler and Lee 2004; Lichtenstein, Drumwright, and Bridgette 2004). According to a recent special report in *BusinessWeek* (Berner 2005, p. 72), large companies disclosed substantial investments in CSR initiatives (i.e., Target's donation of $107.8 million in CSR represents 3.6% of its pretax profits, General Motors's donation of $51.2 million represents 2.7% of its pretax profits, General Mills's donation of $60.3 million represents 3.2% of its pretax profits, Merck's donation of $921 million represents 11.3% of its pretax profits, and Hospital Corporation of America's donation of $926 million represents 43.3% of its pretax profits). By dedicating ever-increasing amounts to cash donations, in-kind contributions, cause marketing, and employee volunteerism programs, companies are acting on the premise that CSR is not merely the "right thing to do" but also "the smart thing to do" (Smith 2003, p. 52).

Importantly, along with increasing media coverage of CSR issues, companies themselves are also taking direct and visible steps to communicate their CSR initiatives to various stakeholders, including consumers. A decade ago, Drumwright (1996) observed that advertising with a social dimension was on the rise. The trend seems to continue. Many companies, including the likes of Target and Wal-Mart, have funded large national ad campaigns promoting their good works. The October 2005 issue of *InStyle* magazine alone carried more than 25 "cause" advertisements. Indeed, consumers seem to be taking notice; whereas in 1993, only 26% of people surveyed by Cone Communications could name a company as a strong corporate citizen, by 2004, the percentage surged to as high as 80% (Berner 2005).

Motivated, in part, by this mounting importance of CSR in practice, several marketing studies have found that social responsibility programs have a significant influence on several customer-related outcomes (Bhattacharya and Sen 2004). More specifically, on the basis of lab experiments, CSR is reported to affect, either directly or indirectly, consumer product responses (Brown 1998; Brown and Dacin 1997), customer–company identification (Sen and Bhattacharya 2001), customer donations to nonprofit organizations (Lichtenstein, Drumwright, and Bridgette 2004), and, more recently, customers' product attitude (Berens, Van Riel, and Van Bruggen 2005).

Although this stream of research has contributed a great deal of insight, there is still a limited understanding of whether and how CSR affects financial outcomes of the firm, such as its market value. Yet it is important to evaluate CSR's impact on market value (i.e., stock-based firm performance) because a firm's financial health is the ultimate test for the success or failure of any strategic initiative. Moreover, prior laboratory studies and anecdotal examples are yet to be complemented with a large-scale analysis using secondary data. Indeed, Brown and Dacin (1997, p. 80) urgently call for research on "how societally oriented activities might bring about positive outcomes for the firm." Echoing this, Berens, Van Riel, and Van Bruggen (2005)

Xueming Luo is Assistant Professor of Marketing, Department of Marketing, University of Texas at Arlington (e-mail: luoxm@uta.edu). C.B. Bhattacharya is Associate Professor of Marketing, Department of Marketing, School of Management, Boston University (e-mail: cb@bu.edu). The authors thank Biao He, Khurram Ansari, Thitikarn Rasrivisuth, and Xiaochu Yu for their assistance with data collection and analysis. They also thank the anonymous *JM* reviewers, Guido Berens, Donald Lichtenstein, Fernado Jaramillo, and seminar participants at the University of Texas at Arlington for their constructive and insightful comments on previous versions of this article.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

energetically call for research efforts that directly link CSR to stock market performance.

Our research responds to this call by investigating the linkage between CSR and firm market value with a longitudinal, archival data set. In keeping with contingent linkages between CSR and consumer responses that prior researchers articulated (see, e.g., Bhattacharya and Sen 2004), we do not predict a simple, unconditional relationship between CSR and market value. This is because firms are not the same in executing, supporting, and exploiting CSR initiatives in the marketplace (Brown 1998; Sen and Bhattacharya 2001). Specifically, companies may generate different (i.e., positive, nonsignificant, and negative) market returns from CSR under different conditions. For example, Starbucks's superior brand equity and its successful CSR initiatives with the charity agency CARE are due, at least in part, to its superior product quality, innovative skills, and ability to obtain and sustain customer satisfaction over time. In contrast, many companies find that CSR results in negative financial returns because of the added costs of making extensive charitable contributions and the diverted attention from improving product quality that would have allowed them to better satisfy customer needs and wants (McGuire, Sundgren, and Schneeweis 1988; Sen and Bhattacharya 2001). Thus, the research questions in this study are as follows: (1) Under what conditions do CSR initiatives result in positive financial performance? and (2) Does customer satisfaction matter in the relationship between CSR and firm performance?

To address these questions, we develop and test a conceptual model that proposes that CSR initiatives enable firms to build a base of satisfied customers, which in turn contributes positively to market value. Specifically, we predict that customer satisfaction partially mediates the relationship between CSR and market value. Although extant marketing literature has addressed the direct impact of customer satisfaction on firm shareholder value (e.g., Anderson, Fornell, and Mazvancheryl 2004; Fornell et al. 2006), the mediating role of customer satisfaction in the financial contribution of CSR has been ignored. In this study, we explicitly theorize this role and argue that building customer satisfaction represents part of the underlying mechanism through which the financial promises of CSR are capitalized.

Furthermore, we explore the boundary conditions under which firms may derive positive or negative market value from CSR. Drawing on various theoretical bases, we argue that firms that have better inside-out corporate abilities (i.e., product quality and innovativeness) to begin with tend to generate more market value from outside-in strategic initiatives (i.e., CSR programs). Conversely, firms that exhibit poorer corporate abilities may find that CSR actually harms customer satisfaction and, because of the lowered satisfaction, decreases stock performance.

Based on multiple secondary data sets that comprise ratings of large companies, the results show support for the CSR → customer satisfaction → firm market value causal linkage. In addition, we find that a proper combination of external CSR initiatives and internal corporate abilities can lead to synergistic returns. However, the data also reveal a previously neglected "dark side" of CSR. That is, CSR actually reduces customer satisfaction levels in firms with low innovativeness capability and, through this negative impact, harms firm market value. The uncovered mediated and asymmetrically moderated results suggest a more nuanced understanding of the financial returns to CSR for both practitioners and marketing researchers.

Conceptual Framework and Hypotheses

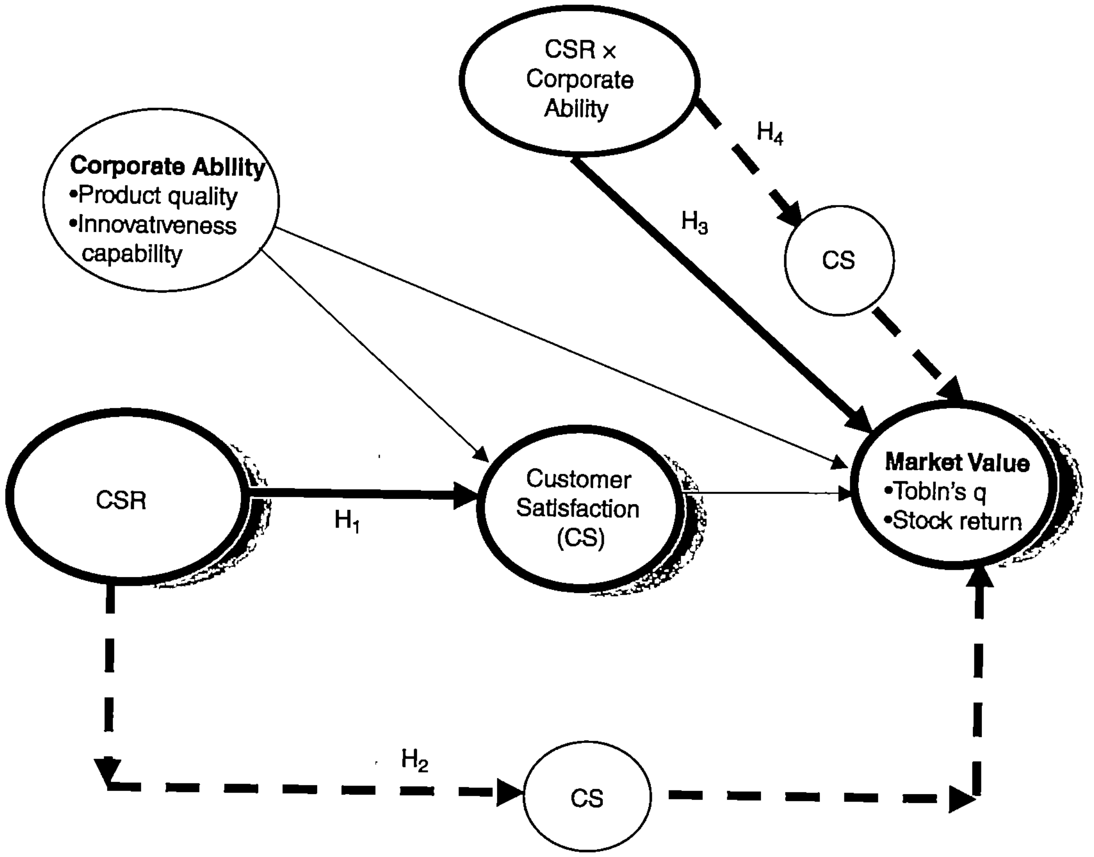
CSR and Market Value

Broadly defined, CSR is a company's activities and status related to its perceived societal or stakeholder obligations (Brown and Dacin 1997; Sen and Bhattacharya 2001; Varadarajan and Menon 1988). Although studies in strategy and finance have explored the relationship between CSR actions and firm performance, empirical evidence to date has been rather conflicting (for a review, see Orlitzky, Schmidt, and Rynes 2003; Pava and Krausz 1996). For example, the returns to CSR are found to be positive in some studies (e.g., Fombrun and Shanley 1990; Soloman and Hansen 1985) but negative in others (e.g., Aupperle, Carroll, and Hatfield 1985; McGuire, Sundgren, and Schneeweis 1988). Thus, Margolis and Walsh (2003, p. 277) conclude that the relationships between CSR and financial performance are decisively "mixed."

There are at least two explanations for these conflicting findings. First, existing studies have largely related CSR to backward-looking firm profitability (i.e., accounting-based return on investment) but not to forward-looking firm market value (i.e., stock-based Tobin's q). Theoretically, however, market value is different from (and perhaps more important than) return on investment because "accounting measures are retrospective and examine historical performance. In contrast, the market value of firms hinges on growth prospects and sustainability of profits, or the expected performance in the future" (Rust, Lemon, and Zeithaml 2004, p. 79). Second, the equivocal link between CSR and firm performance may be due, in part, to extant strategy and finance literature having largely omitted the underlying processes or contingency conditions that may explain the range of observed relationships (Sen and Bhattacharya 2001).

We precisely examine these research issues in this study. In particular, as we show in Figure 1, our framework proposes that the relationship between CSR and firm market value is better understood by the mediating link of customer satisfaction. In recent times, scholars (e.g., Anderson, Fornell, and Mazvancheryl 2004; Fornell et al. 2006) have demonstrated the positive relationship between customer satisfaction and market value. We build on this literature and institutional theory to propose that CSR is a driver of customer satisfaction and that the CSR–firm market value linkage exists (at least partially) because of the underlying process through customer satisfaction. In addition, drawing on work in the area of corporate identity and associations (e.g., Brown and Dacin 1997), we posit that a firm's corporate abilities (i.e., product quality and innovativeness capa-

FIGURE 1
Conceptual Framework



Notes: Bolded paths are hypothesized relationships. Unbolded paths have been studied previously (e.g., Anderson, Fornell, and Mazvancheryl 2004; Anderson, Fornell, and Rust 1997; Fornell et al. 2006; Griffin and Hauser 1996; Mithas, Krishnan, and Fornell 2005b). Dashed paths indicate that the depicted relationships are partially mediated by customer satisfaction.

bility) moderate the relationship between CSR and market value. Finally, we expect that customer satisfaction mediates, at least partially, these moderated relationships.

CSR and Customer Satisfaction

Customer satisfaction is defined as an overall evaluation based on the customer's total purchase and consumption experience with a good or service over time (Anderson, Fornell, and Mazvancheryl 2004; Fornell 1992). In the marketing literature, customer satisfaction has been recognized as an important part of corporate strategy (Fornell et al. 2006) and a key driver of firm long-term profitability and market value (Gruca and Rego 2005).

Why should a firm's CSR initiatives lead to greater customer satisfaction? At least three research streams point to such a link: First, both institutional theory (Scott 1987) and stakeholder theory (Maignan, Ferrell, and Ferrell 2005) suggest that a company's actions appeal to the multidimensionality of the consumer as not only an economic being but also a member of a family, community, and country (Handelman and Arnold 1999). Building on this, Daub and Ergenzinger (2005) propose the term "generalized customer" to denote people who are not only customers who care about the consumption experience but also actual or potential members of various stakeholder groups that companies need to consider. Viewed in this way, such generalized customers are likely to be more satisfied by products and services that socially responsible firms (versus socially irresponsible counterparts) offer.

Second, a strong record of CSR creates a favorable context that positively boosts consumers' evaluations of and attitude toward the firm (Brown and Dacin 1997; Gürhan-Canli and Batra 2004; Sen and Bhattacharya 2001). Specifically, recent works on customer–company identification (Bhattacharya and Sen 2003, 2004) suggest that CSR initiatives constitute a key element of corporate identity that can induce customers to identify (i.e., develop a sense of connection) with the company. Indeed, Lichtenstein, Drumwright, and Bridgette (2004, p. 17) note that "a way that CSR initiatives create benefits for companies appears to be by increasing consumers' identification with the corporation … [and] support for the company." Not surprisingly, identified customers are more likely to be satisfied with a firm's offerings (e.g., Bhattacharya, Rao, and Glynn 1995; Bhattacharya and Sen 2003).

The third literature stream that enables us to relate CSR to customer satisfaction examines the antecedents of customer satisfaction. For example, perceived value is a key antecedent that has been empirically shown to promote customer satisfaction (Fornell et al. 1996; Mithas, Krishnan, and Fornell 2005b). In our context, all else being equal, customers likely derive better perceived value and, consequently, higher satisfaction from a product that is made by a socially responsible company (i.e., added value through good social causes). Furthermore, engaging in CSR may allow firms to understand their generalized customers better and thus improve their customer-specific knowledge (Sen and Bhattacharya 2001). Because improving customer knowledge represents another antecedent that has been found to enhance customer satisfaction (Jayachandran et al. 2005; Mithas, Krishnan, and Fornell 2005a), we believe that CSR initiatives may help promote customer satisfaction.

> H_1: All else being equal, firms that are viewed more favorably for their CSR initiatives enjoy greater customer satisfaction.

The Mediating Role of Customer Satisfaction

The existing marketing literature shows accumulating evidence for the influence of customer satisfaction on firm market value. For example, firms with satisfied customers tend to enjoy greater customer loyalty (e.g., Bolton and Drew 1991; Oliver 1980), positive word of mouth (Szymanski and Henard 2001), and customer's willingness to pay premium prices (Homburg, Koschate, and Hoyer 2005), all of which can increase a firm's market value. Indeed, several studies find that firms with higher levels of customer satisfaction are able to achieve higher levels of cash flows (e.g., Gruca and Rego 2005; Fornell 1992; Mittal et al. 2005) and less volatility of future cash flows, thus leading to superior market value (e.g., Anderson, Fornell, and Mazvancheryl 2004; Fornell et al. 2006; Srivastava, Shervani, and Fahey 1998).

In linking this evidence for the influence of customer satisfaction on firm market value with our first hypothesis on the influence of CSR on satisfaction, a mediating role of customer satisfaction in the CSR–performance linkage might logically be expected. That is, CSR affects customer satisfaction, which in turn affects market value. In other words, customer satisfaction represents the mediational pathway through which CSR actions affect firm market value.

However, there may be "noncustomer routes" by which CSR affects market value. For example, both textbooks (e.g., Kotler and Lee 2004; Pava and Krausz 1996) and academic articles (e.g., Godfrey 2005; Margolis and Walsh 2003) have pointed to the impact of CSR on multiple stakeholders, such as employees and investors as well as consumers. In particular, positive "moral capital" as a result of CSR (Godfrey 2005, p. 777) could directly affect market value by improving employee morale and productivity. In addition, CSR creates public goodwill (Houston and Johnson 2000; McGuire, Sundgren, and Schneeweis 1988), which provides an "insurance-like" protection to shareholder wealth. As a consequence, putting the pieces together, we predict a partially mediating role of customer satisfaction on the impact of CSR on market value.

> H_2: All else being equal, firms that are viewed more favorably for their CSR initiatives enjoy higher market value, and a firm's customer satisfaction level at least partially mediates this influence of CSR on market value.

The Moderating Role of Corporate Abilities

In this section, we argue that the relationship between CSR and firm market value may not be universally positive but rather contingent on several boundary conditions. That is, a positive or negative relationship may be observed, depending on the levels of corporate abilities. In general, corporate abilities refer to various elements of a firm's expertise and competency, such as the ability to improve the quality of existing products/services and the ability to generate new products/services innovatively (Gatignon and Xuereb 1997; Rust, Moorman, and Dickson 2002; Zeithaml 2000). According to Brown and Dacin (1997), a company's CSR and corporate abilities *both* influence customers' perceptions of the company's products.

We expect that firms with low levels of corporate abilities (i.e., low levels of innovativeness capabilities and product quality) generate negative market value from CSR for several reasons. On the basis of institutional theory, Handelman and Arnold (1999) contend that companies should engage in CSR with good causes (for the *social* aspect of legitimation) and, at the same time, provide a good product (for the *pragmatic* aspect of legitimation). Thus, it is likely that CSR initiatives fail to generate a favorable impact if the firm is perceived as less innovative and as offering poor-quality products (i.e., due to a lack of pragmatic legitimation; see DiMaggio and Powell 1983). Indeed, Sen and Bhattacharya (2001) show that CSR initiatives may even backfire with reduced purchase intent and negative perceptions if consumers believe that CSR investments are at the expense of developing corporate abilities, such as product quality and innovativeness (i.e., investments represent "misguided priorities" on the part of the firm with low levels of corporate abilities). More important, consumers may make negative and detrimental attributions regarding a firm's motives if a low-innovativeness or low-product-quality firm engages in social responsibility. This would ultimately result in an unattractive corporate identity and, thus, negative market returns by virtue of negative word of mouth and detrimental customer complaints (Brown 1998; Varadarajan and Menon 1988).[1]

Conversely, we predict that firms with high levels of corporate abilities generate positive market value from CSR. Such firms tend to posses better corporate image and more attractive identities with which consumers want to identify (Bhattacharya and Sen 2003). When coupled with high corporate abilities, a firm's CSR actions are more likely to generate favorable attributions and consumer identification. This would ultimately promote performance-enhancing behaviors, such as customer loyalty (Bhattacharya and Sen 2004). Indeed, if a firm can accommodate customers and other stakeholders and meet different sets of norms (e.g., pragmatic and social norms) by not merely executing CSR initiatives but also developing strong corporate abilities to support and exploit these CSR actions, it is in a better position to win the social contract, institutional allegiance, moral legitimacy, and consumers' support for the organization (cf. Handelman and Arnold 1999, p. 34; Scott 1987). Taken together, these beneficial effects suggest a positive market return to CSR for firms with high levels of corporate abilities. Therefore, we propose an asymmetric moderating effect of corporate abilities on the association between CSR and firm market value.

---

[1] Our focus here is on the moderating role of corporate ability in the CSR–performance link rather than on the direct relationship between CSR and corporate ability. That is, we do not investigate whether CSR directly affects or is related to innovativeness and product quality (i.e., corporate ability–related constructs) given the conflicting literature. On the one hand, Brown and Dacin (1997, p. 68, emphasis added) contend that "CSR associations are often *unrelated* to the company's abilities in producing goods and services." On the other hand, it is possible that a firm's innovation is CSR oriented (e.g., environmentally responsible packaging), and CSR initiatives may affect product-quality perceptions.

> H_3: Corporate abilities (i.e., product quality and innovativeness capability) moderate the relationship between CSR and market value. The relationship will be negative for firms with low corporate abilities but will be positive for firms with high corporate abilities.

The Mediating Role of Customer Satisfaction in the Moderated Relationships

Finally, as we have argued, part of the mechanism by which CSR actions influence a firm's market value is customer satisfaction. Thus, it is conceivable that the positive impact of CSR on firms with high levels of corporate abilities enhances the level of customer satisfaction, which then leads to enhanced market value (Anderson, Fornell, and Mazvancheryl 2004; Brown and Dacin 1997; Sen and Bhattacharya 2001).

On the contrary, for firms that are low in corporate ability (i.e., they are neither innovative nor competent in product quality), CSR actions may not be able to generate much institutional legitimacy, customer–company identification, or customer satisfaction (Scott 1987). As a result, CSR initiatives may relate little to financial results and market value (e.g., Margolis and Walsh 2003; Mithas, Krishnan, and Fornell 2005b) in firms with low levels of corporate abilities. Thus:

> H_4: A firm's customer satisfaction at least partially mediates the moderated relationship among CSR, corporate abilities (i.e., product quality and innovativeness capability), and market value.

Data and Variable Construction

In this section, we describe the secondary data that we collected to test the hypotheses. We also present the construction of the variables, such as CSR, corporate abilities, customer satisfaction, and market value. In Table 1, we report the variables, their definitions, and data sources. We collected data for the publicly traded *Fortune* 500 companies from multiple archival sources: COMPUSTAT, *Fortune* America's Most Admired Corporations (FAMA), the American Customer Satisfaction Index (ACSI), Competitive Media Reporting (CMR), and Center for Research in Security Prices (CRSP).

Measuring CSR

One approach to measuring market perceptions of firms' CSR initiatives is to rely on the amount of CSR investments disclosed in firms' annual reports to shareholders. However, there are many important doubts about the validity of the announced CSR investments, despite the seeming attractiveness of this approach. For example, there is a lack of consensus on what should be included (or excluded) in CSR investments (Margolis and Walsh 2003; Orlitzky, Schmidt, and Rynes 2003; Tsoutsoura 2004). Few companies have their announced CSR investments audited or validated externally by third parties. Thus, some firms may overreport CSR investments for impression management (i.e., exaggerating their giving). Other firms may underreport CSR investments because they may regard CSR investments only as donated cash or in-kind products and services (excluding investments that benefit the environment and their employees). Furthermore, although some external sources (e.g., 100 best corporate citizens by *Business Ethics*, csrwire.com, Social Responsibility Initiative reports) may track companies' CSR investments objectively, the nature and amount of CSR investments for the same firm can change dramatically from one source to another (Berner 2005; Fombrum and Shanley 1990; Margolis and Walsh 2003).

Therefore, we turn to subjective measures of CSR. Although some studies use small-scale survey data with a limited set of firms (e.g., Christmann 2000), prior research suggests the use of a more comprehensive, large-scale survey data set available from FAMA to measure CSR (McGuire, Sundgren, and Schneeweis 1988). More specifically, in ranking the United States' most admired corporations each year, FAMA polls more than 10,000 financial analysts, senior executives, and Wall Street investors from more than 580 large companies (see *Fortune* 2005, p. 68).

For each firm-year observation, FAMA collects ratings of CSR that have been made on an interval scale ranging from 0 to 10, with 10 as the highest; the ratings represent a comparison among major competing companies in a given industry. Studies in both marketing and strategy (e.g., Fombrum and Shanley 1990; Houston and Johnson 2000; McGuire, Sundgren, and Schneeweis 1988) have reported evidence of reliability and validity of this data source. In particular, McGuire, Schneeweis, and Branch (1990, p. 170) note, "*Fortune* reputation is one of the most comprehensive and widely circulated surveys of attributes available. Both the quality and number of respondents are comparable or superior to the 'expert panels' usually gathered for such purposes." Houston and Johnson (2000, p. 12) also acknowledge it as the "best secondary" data source.

Prior research has shown that there is a reverse-causality concern between CSR and financial performance (e.g., McGuire, Sundgren, and Schneeweis 1988). That is, a firm's CSR affects its future performance, and a firm's history of financial performance contributes to its current CSR involvement. We accommodate this concern by using the approach that Roberts and Dowling (2002) recommend. In

TABLE 1
Variables and Data Sources

Variables	Definitions; Measures	Secondary Data Sources	Data Types
CSR	Broadly defined as a company's activities and status related to its perceived societal or stakeholder obligations; *latent* variable indicated by CSR scores in 2001 (published in 2002), 2002 (published in 2003), and 2003 (published in 2004).	FAMA	Interval from 0 to 10
Customer satisfaction	Defined as an overall evaluation of the postconsumption experience of products or services in the minds of customers; *latent* variable indicated by customer satisfaction scores in 2002, 2003, and 2004.	ACSI	Interval from 0 to 100
Product quality	Defined as the minimum condition or the threshold of product attributes that a firm must meet when offering its products or service in competitive markets; *latent* variable indicated by quality of products/services scores in 2001, 2002, and 2003.	FAMA	Interval from 0 to 10
Innovativeness capability	Defined as a firm's ability to apply its internal knowledge stock to produce new technology, new products/services, and other new fronts; *latent* variable indicated by quality of products/services scores in 2001, 2002, and 2003.	FAMA	Interval from 0 to 10
Tobin's q	Stock price–based measure of firm market value; *observed* variable based on the average of Tobin's q in 2002, 2003, and 2004.	CRSP COMPUSTAT	Ratio
Stock return	Stock price–based measure of firm market value; *observed* variable based on the average of stock return in 2002, 2003, and 2004.	CRSP COMPUSTAT	Ratio

particular, we regress CSR scores against firm financial performance (return on assets [ROA]) in the prior four years and save the residual of this regression as the final measure of CSR. Because this residual is independent from financial performance, the reverse-causality bias is no longer a concern.

Following the work of Cho and Pucik (2005), we used the ratings of CSR for each firm in 2001, 2002, and 2003 (but published in 2002, 2003, and 2004, due to a one-year lag in print) as three separate indicators of the latent construct of CSR.[2] This approach of using measurement items with different time frames is also widely applied in the strategy (e.g., Li and Atuahene-Gima 2001) and personal-selling and psychology literature streams (Bluedorn 1982; Boles, Johnston, and Hair 1997; Johnston et al. 1990; Netemeyer, Maxham, and Pullig 2005).

Measuring Corporate Abilities

We do not view corporate abilities simply as a unidimensional construct. Instead, we consider two specific corporate abilities: product quality and innovativeness capability (Gatignon and Xuereb 1997; Rust, Moorman, and Dickson 2002; Zeithaml 2000). In our view, both innovativeness and product quality can represent the dimensions of corporate ability that Brown and Dacin (1997) propose. Although product quality refers to a firm's ability to "exploit" the capabilities of products already in the marketplace (Cho and Pucik 2005; March 1991), innovativeness represents a firm's ability to "explore" new market possibilities in terms of developing new products (Kim and Mauborgne 1997; Kleinschmidt and Cooper 1991). In addition, commitment to the quality of existing products is essential for keeping a firm's current customers happy, whereas innovation is essential for reaching new customer bases and catering to ever-changing customer needs.

---

[2] Cho and Pucik (2005) find strong support (construct and criterion-related validity) for using multiyear ratings from *Fortune* magazine as indicators of the underlying latent variable McGuire, Sundgren, and Schneeweis (1988) also employ single-year CSR ratings from *Fortune* as the measure of CSR.

Formally, product quality can be defined as the minimum condition or the threshold of product attributes that a firm must meet when offering its products/services in competitive markets (Rust, Moorman, and Dickson 2002; Vargo and Lusch 2004; Zeithaml, Parasuraman, and Berry 1990). Prior studies have established that a firm's ability to provide a superior product/service quality is critical for its long-term survival and success (e.g., Buzzell, Gale, and Sultan 1975; Mittal et al. 2005; Rust, Moorman, and Dickson 2002).

In a similar fashion to CSR, we measure product quality by FAMA ratings in 2001, 2002, and 2003 (published in 2002, 2003, and 2004) as the underlying indicators. Again, because of the reverse causality between financial performance and FAMA ratings, we control for this bias and obtain clean measures for product quality and innovativeness capability by employing the same residual approach as in the case of CSR (e.g., Roberts and Dowling 2002).

Innovativeness capability is a firm's ability to apply its internal knowledge stock to produce new technology, new products/services, and other new fronts (Drucker 1993; Griffin and Hauser 1996). According to exploration learning theory (March 1991), innovation is also critical for the survival and success of organizations because dynamic markets constantly shake out the players that lack capabilities to explore new market opportunities (Gatignon and Xuereb 1997; Schumpeter 1934). Similar to product quality, we measure the latent variable of a firm's innovativeness capability by using its *Fortune* ratings in 2001, 2002, and 2003 from FAMA (published in 2002, 2003, and 2004) as three separate indicators underlying this construct. Prior research has employed this data source to measure companies' innovativeness capability (Cho and Pucik 2005).

Measuring Customer Satisfaction

We used the ACSI database to measure customer satisfaction. In the marketing literature, the ACSI has been shown to be a reliable source of measuring customer satisfaction. Several studies employ this database to assess overall customer satisfaction of total purchase and consumption experience at the firm level (e.g., Anderson, Fornell, and Mazvancheryl 2004; Fornell et al. 2006; Gruca and Rego 2005; Mithas, Krishnan, and Fornell 2005b; Mittal et al. 2005). The National Quality Research Center at the University of Michigan developed and maintains the ACSI data set. It has data for nearly 200 *Fortune* large companies that span all major economic sectors and constitute approximately 43% of the U.S. economy. To obtain ACSI data, more than 50,000 household consumers (actual product users) of these large firms are polled on a quarterly basis. Each valid respondent has passed screening questions related to predefined purchase and consumption periods. The ACSI uses an interval scale ranging from 0 to 100, with 100 as the highest level of customer satisfaction.

Based on multi-item, multiconstruct criteria, the ACSI is a reliable data source because it employs the same survey questionnaire, random sampling, and estimation modeling across firms and years (Fornell et al. 1996; Fornell et al. 2006; Mithas, Krishnan, and Fornell 2005b). A comprehensive test of the validity and reliability of this satisfaction measure can be found in the work of Fornell and colleagues (1996). Parallel to CSR, innovativeness capability, and product quality, we treat customer satisfaction as a latent variable and measure it using its ACSI ratings in 2002, 2003, and 2004 as three separate indicators.

Measuring Market Value

We have two separate measures of market value at the firm level across years: Tobin's q and stock return. We follow prior marketing studies (Lee and Grewal 2004; Rao, Agarwal, and Dahlhoff 2004) to calculate Tobin's q for each firm-year observation.[3] In addition, following Jacobson and colleagues (i.e., Aaker and Jacobson 1994, 2001; Mizik and Jacobson 2003), we derive the measure of stock return using the COMPUSTAT and CRSP databases.[4] Rather than using a simple year-end stock price, we use a more conservative measure of stock price—that is, the average of the end of the four quarters of stock prices—when calculating Tobin's q and stock return (Lee and Grewal 2004). We then use the derived three-year average (2002, 2003, and 2004) of Tobin's q and stock return as observed measures for market value. Compared with market value, the predicting variables of CSR, innovativeness capability, and product quality were all lagged by one year to be more precise on the specific direction of causality and to reduce the possibility of endogeneity bias (Murthi, Srinivasan, and Kalyanaram 1996; Rust, Moorman, an Dickson 2002).

Measuring Control Variables

We obtained the data for control variables such as research-and-development (R&D) intensity, firm size, competition intensity, and ROA from COMPUSTAT, and we obtained the data for advertising intensity from CMR. More specifically, R&D intensity is the ratio of R&D spending to total assets. We control for the influence of R&D expenditures on performance because a firm's R&D intensity enhances innovation activities and investors' evaluations of the firm (Chauvin and Hirschey 1993; Gruca and Rego 2005; McGuire, Sundgren, and Schneeweis 1988).

Advertising intensity is the ratio of reported advertising spending to total assets. Because COMPUSTAT has many missing data points for firm advertising expenditure, we use the CMR database for advertising-spending data (Rao, Agarwal, and Dahlhoff 2004). We control for the influence of advertising expenditures on performance because intense advertising promotes customer awareness, brand equity, and sales revenues (e.g., Joseph and Richardson 2002; Morgan and Rego 2006).

---

[3] Rao, Agarwal, and Dahlhoff (2004, p. 130) provide a detailed function on how to derive Tobin's q. That is, q = (share price × number of common stock outstanding + liquidating value of the firm's preferred stock + short-term liabilities – short-term assets + book value of long-term debt)/book value of total assets.

[4] In particular, Aaker and Jacobson (2001, p. 489) and Mizik and Jacobson (2003, p. 71) suggest a detailed function on how to calculate stock return. That is, stock return = (current year's share price × number of common stock outstanding + dividends – previous year's share price × number of common stock outstanding)/(previous year's share price × number of common stock outstanding).

Firm size is the log of number of employees. We control for the influence of firm size because large firms may have more resources and thus enjoy economies of scale, but small firms may have higher strategic flexibility when seeking entrepreneurial rents (Dutta, Narasimhan, and Rajiv 1999; Rao, Agarwal, and Dahlhoff 2004).

Strategic focus is the number of business segments in which the firm operates (Rao, Agarwal, and Dahlhoff 2004). This variable is available directly from the menu choice at the Compact Disclosure (CD-ROM), which defines it as "the number of unique business segments of an individual company." We control for this influence because of possible diversification effects. That is, more diversified firms may have a faster asset turnover rate and exhibit economies of scope. However, highly diversified firms may lack focus in the highly segmented, competitive marketplace and thus experience negative returns (Fombrum and Shanley 1990; Gruca and Rego 2005).

We measure competition intensity by using the Herfindahl concentration index, derived from COMPUSTAT. Following prior work (Gruca and Rego 2005; Mithas, Krishnan, and Fornell 2005a), we calculate this concentration index at the primary four-digit industry level of Standard Industrial Classification codes (which has been replaced by the North American Industry Classification System) for each firm-year observation. We use this covariate to control for impact of industry competition level (Rao, Agarwal, and Dahlhoff 2004).

Finally, we control for the influence of ROA in predicting stock return and Tobin's q (Chauvin and Hirschey 1993). In particular, we measure ROA as the ratio of net income after extraordinary items to book value of total assets, derived from COMPUSTAT. We used the average of the 2002, 2003, and 2004 data points as the measure of ROA. We include ROA as a covariate variable because of the impact of financial information on the stock market (Chauvin and Hirschey 1993; Erickson and Jacobson 1992). Table 2 presents the summary statistics for all variables in this study.

Despite having these stringent controls, in light of our moderation hypotheses, a lingering issue is whether there are systematic industry differences between firms that are rated high on product quality (and/or innovativeness) and those that are rated low. A close examination of the top and bottom firms on the dimensions of product quality and innovativeness allays this concern. We find that both the top and the bottom firms in terms of their innovativeness and product quality ratings cover a variety of industries, such as retail, services, and manufacturing. More specifically, top innovativeness firms include Apple, Google, Procter & Gamble, FedEx, Nike, and Target, among others; bottom innovativeness firms include United Airlines, Dillard's, Kmart, and Qwest Communications, among others, according to *Fortune*'s large-scale survey data in 2005. In other words, neither the low-innovativeness nor the high-innovativeness firms are dominated by particular industry types.

Merged Final Data Set

We merged data from these different archival sources and obtained unbalanced panel time-series, cross-sectional data consisting of 452 firm-year observations across 113 firms for the 2001–2004 periods. However, one year's data are lost because we employed the lagging process (2001–2003 for CSR, product quality, and innovativeness; 2002–2004 for customer satisfaction, Tobin's q, and stock return) to reduce the endogeneity bias and reverse-causality concerns described previously. Thus, we were able to use 339 data points for hypotheses testing. This merged data set includes individual firms in various industries, ranging from durables (e.g., automobiles, household appliances, personal computers), to nondurables (e.g., cigarettes; athletic shoes; services, such as airlines, hotels, and utilities), to retail (e.g., department stores, discount stores, supermarkets), among others. Although FAMA has ratings of CSR, innovativeness capability, and product quality for approximately 580 firms (Cho and Pucik 2005; *Fortune* 2005) and ACSI has data on approximately 190 firms/brands (Fornell et al. 1996; Fornell et al. 2006; Gruca and Rego 2005; Morgan and Rego 2006), we were not able to obtain a larger sample of firms in the merged final data set. This is because many firms included in *Fortune*'s source are not represented in the ACSI source and because the same firm may have several brands in the ACSI (Anderson, Fornell, and Mazvancheryl 2004). We also tried to search other relevant secondary sources (Standard & Poor's industry reports, company annual reports, Compact Disclosure, and Moody's report) to cross-validate our final data set spanning the period from 2001 to 2004.

Note that COMPUSTAT does not have complete data points for all variables. For example, because COMPUSTAT does not require companies to report their R&D investments (volunteered responses only; see Joseph and Richardson 2002), we found that more than 40% of observations for the control variable of R&D are missing across the years. Before testing the hypotheses, we controlled for the covariates using the same approach applied in prior studies (e.g., Ahearne, Bhattacharya, and Gruen 2005; Pan, Ratchford, and Shankar 2002). In particular, we ran a linear regression with all control variables (firm-level and industry-level) as independent variables and Tobin's q as the dependent variable. We saved the unstandardized residuals from this regression and then used them as the surrogate for Tobin's q in all structural equation models (SEMs). We also applied this approach to obtain the surrogates for stock return.

Analyses and Results

Measurement Model Results

Following the work of Anderson and Gerbing (1988), we employ confirmatory factor analysis (CFA) to test the validity of the measures. Overall model statistics show that the chi-square for the model is 90.73 (d.f. = 48, $p > .05$),

TABLE 2
Descriptive Statistics of Variables

Variables	M	SD	1	2	3	4	5	6	7	8	9	10	11	12
1. CSR	6.03	.95	1.00											
2. Customer satisfaction	79.06	8.77	.20**	1.00										
3. Innovativeness capability	6.06	1.42	.72**	.18**	1.00									
4. Product quality	5.97	1.10	.78**	.22**	.81**	1.00								
5. Tobin's q	1.82	1.35	.13*	.18**	.13*	.10	1.00							
6. Stock return	.29	.46	.14*	.17**	.11	.14*	.42**	1.00						
7. Advertising intensity	.06	.08	.06	.07	.01	.02	.11*	.08	1.00					
8. R&D intensity	.04	.07	.03	.05	.04	.02	.10	.06	.04	1.00				
9. Firm size (log of asset)	4.16	.90	.05	.02	.04	.05	.09	.07	.11	.09	1.00			
10. Strategic focus	2.87	1.68	–.03	–.03	–.02	–.00	–.07	–.03	–.03	–.02	–.04	1.00		
11 Competition intensity	.07	.13	–.04	.02	–.04	–.05	–.08	–.05	–.04	–.05	–.03	.01	1.00	
12. ROA	3.68	9.87	.19**	.22**	.20**	.19**	.33**	.28**	.10	.06	.08	–.03	–.02	1.00

*$p <$ 05.
**$p < .01$.

and the comparative fit index (CFI), goodness-of-fit index (GFI), and root mean square error of approximation (RMSEA) are satisfactory (.94, .92, and .06, respectively).

As we report in Table 3, the CFA results lend some support for the convergent validity for all the measures because all estimated loadings of indicators for the underlying constructs are significant (i.e., smallest t-value = 6.53, $p < .05$). Cronbach's alpha of the constructs exceeded the .7 threshold (Nunnally 1978). The minimum reliability of these measures is .85, as we reported. In addition, the average variance extracted (AVE) across the constructs exceeds the .5 benchmark (see Fornell and Larcker 1981). As Table 3 shows, the smallest AVE of the constructs is .72. The data also supported discriminant validity of the measures. We examined pairs of measures using the constrained model and unconstrained model in a series of chi-square difference tests (Anderson and Gerbing 1988). The test results consistently indicated that for each pair of constructs, the unconstrained models fit the data better than their constrained counterparts, suggesting discriminant validity. In addition, we compared the estimated AVE of each measure with the squared correlation between-measure pairs (Fornell and Larcker 1981). In all cases, we found that the AVEs exceeded the squared correlations, further confirming the discriminant validity of the constructs.

Results for the Mediating Role of Customer Satisfaction

In testing the mediating role of customer satisfaction, we used SEM to consider explicitly the possible bias of measurement error on path estimates. Consistent with the procedures in psychology (e.g., Holmbeck 1997) and marketing (Andrews et al. 2004; Handelman and Arnold 1999; Selnes and Sallis 2003), our SEMs not only account for measurement error but also allow for a comprehensive test of the hypotheses related to mediation, moderation, and mediated moderation.

Table 4 reports the results of the SEMs. H_1 predicted that CSR would positively affect customer satisfaction. Model 1 examines this prediction, and the result is statistically significant, in support of H_1. We assess the significance of the reported SEM path estimates through a bootstrapping approach with 1000 resamples. As the CFI, GFI, and RMSEA indicate, Model 1 fits the data well.

H_2 predicted that CSR would positively influence a firm's market value and that customer satisfaction would mediate this influence. To establish the existence of this mediation effect, four conditions should hold (Andrews et al. 2004): (1) The predictor variable (CSR) should significantly influence the mediator variable (customer satisfaction); (2) the mediator should significantly influence the dependent variable (market value); (3) the predictor (CSR) variable should significantly influence the dependent variable (market value); and (4) after we control for the mediator variable (customer satisfaction), the impact of the predictor (CSR) on the dependent variable (market value) should no longer be significant (for full mediation) or should be reduced in strength (for partial mediation) (Baron and Kenny 1986, p. 1177).

As Table 4 shows, Model 1 meets the first two conditions. That is, CSR affects customer satisfaction. Furthermore, satisfaction significantly affects both Tobin's q and stock return, which is consistent with existing studies (Anderson, Fornell, and Mazvancheryl 2004; Bolton and Drew 1991; Fornell et al. 2006). Model 2 qualifies the third condition; the predictor variable of CSR affects market value in terms of Tobin's q and stock return. As Table 4 shows, Model 2 does not include the mediator of customer satisfaction and appears to fit the data reasonably well. The fourth condition holds if the effects of CSR on market value become insignificant or less significant after the mediator of customer satisfaction is included. Model 3 results (no-mediation model in Table 4) show that the inclusion of customer satisfaction diminishes the strength of the effect of CSR on firm market value. The main effects of CSR on both Tobin's q and stock return are no longer significant.[5] Thus,

TABLE 3
Results of the CFA

Construct	Items	Factor Loading	t-Value	AVE	CR
CSR				.75	.90
	CSR → CSR01	.69	13.43		
	CSR → CSR02	.71	13.55		
	CSR → CSR03	.75	13.54		
Innovativeness capability (IN)				.74	.87
	IN → IN01	.67	11.10		
	IN → IN02	.68	11.16		
	IN → IN03	.62	9.98		
Product quality (PQ)				.76	.91
	PQ → PQ01	.78	13.39		
	PQ → PQ02	.83	13.92		
	PQ → PQ03	.80	13.51		
Customer satisfaction (CS)				.72	.85
	CS → CS02	.50	7.21		
	CS → CS03	.48	6.92		
	CS → CS04	.46	6.53		

Notes: All t-values are significant ($p < .05$); $\chi^2 = 90.73$ (d.f. = 48, $p > .05$), CFI = .94, GFI = .92, and RMSEA = .06. CR = construct reliability.

---

[5] We also employed ordinary least squares to test the mediation hypotheses. The results are consistent and suggest strong support for the mediation results of CSR. However, because SEM offers at

TABLE 4
SEM Results for Mediation Effects

	SEM Estimates					
Model Specifications	χ^2	**d.f.**	χ^2_{diff} (d.f.$_{diff}$)	**CFI**	**GFI**	**RMSEA**
Model 1	362.10	101	Compared base	.94	92	.05
Model 2	112.82	59		.92	.91	.07
Model 3	391.58	96	29.48** (5)[a]	.91	.89	.07
Model 4	345.05	97	17.05** (4)[b]	.96	.94	.04

	Full Mediation: Model 1	PV → DV: Model 2	Nonmediation: Model 3	Partial Mediation: Model 4
CSR → TQ		.14*	.09	
IN → TQ		.10	.10	.11*
PQ → TQ		.17**	.14*	.12*
CSR × IN → TQ		.14*	.09	
CSR × PQ → TQ		.20**	.15*	.13*
CSR → CS	.23**			.23**
IN → CS	.20**			.19**
PQ → CS	.28**			.26**
CSR × IN → CS	.12*			.14*
CSR × PQ → CS	.18**			.18**
CSR → SR		.13*	.08	
IN → SR		.08	.07	
PQ → SR		.11*	.09	
CSR × IN → SR		.10	.07	
CSR × PQ → SR		.18**	.12*	.11*
CS → TQ	.25**		.23**	.22**
CS → SR	.22**		.21**	.19**
R^2				
CS	.34		.32	
TQ	.46	.41	.45	.48
SR	.38	.34	.37	.39

*$p < .05$, one-tailed test.
**$p < .01$, one-tailed test.
[a]The results of the difference between Model 1 and Model 3.
[b]The results of the difference between Model 1 and Model 4

Notes: CS = customer satisfaction, IN = innovativeness capability, PQ = product quality, TQ = Tobin's q, and SR = stock return. Model 2 (PV → DV) does not include the mediator of customer satisfaction. Model 3 (nonmediation effects) includes the mediator of customer satisfaction.

customer satisfaction seems to mediate fully the direct impact of CSR on firm market value (though it does not mediate fully the interaction effects between CSR and corporate abilities on market value, as we detail next). As such, the data provide strong support for H_2, which predicted that CSR would increase a firm's long-term financial performance through the mediator of customer satisfaction.[6]

Results for the Moderating Role of Corporate Abilities

H_3 predicted that corporate abilities, such as innovativeness capability and product quality, would moderate the impact of CSR on market value. Table 5 reports the hierarchical SEM results related to moderation effects. Following the work of Aiken and West (1991), we mean-centered the CSR, innovativeness capability, and product-quality variables before generating the interaction terms, and then we added the interaction terms hierarchically from Model 2 to Model 3.[7] The results in Table 5 show that the interaction

---

least a weak test of causal pathways and easily compares different rival models, but ordinary least squares does not account for measurement error, we report the results based on the SEMs for all hypotheses in this study.

[6] We also tested the hypotheses with single-year items for the predicting and dependent variables (rather than the reported multiple-years-based separate items). The results are similar in pattern and further support the hypotheses.

[7] Multicollinearity bias was not a severe problem. The highest variance inflation factor was 3.06, and the largest condition index was 3.51. Note that in a mean-centered interaction-effects model, the estimated coefficient of one independent variable is obtained under the assumption of the mean value of other variables. Moreover, the entry of the interactions terms for CSR, innovativeness

TABLE 5
Hierarchical SEM Results

SEM Estimates						
Rival Models	χ^2	**d.f.**	χ^2_{diff} (d.f. $_{diff}$)	**CFI**	**GFI**	**RMSEA**
Model 1	15.83	5	96.99** (54)[a]	.90	.86	.08
Model 2	177.09	40	64.27** (19)[b]	.91	.89	.07
Model 3	112.82	59	Compared base	.92	.91	.07

	Direct Effects: Model 1	Direct Effects: Model 2	Moderated Effects: Model 3
CSR → TQ	.14*	.12*	.14*
IN → TQ		.11*	.10
PQ → TQ		.17**	.17**
CSR × IN → TQ			.14*
CSR × PQ → TQ			.20**
CSR → SR	.12*	.13*	.13*
IN → SR		.08	.08
PQ → SR		.10	.11*
CSR × IN → SR			.10
CSR × PQ → SR			.18**
R^2			
TQ	.30	.35	.41
SR	.28	.29	.34

*$p <$ 05, one-tailed test.
**$p < .01$, one-tailed test.
[a]The results of the difference between Model 1 and Model 3.
[b]The results of the difference between Model 2 and Model 3
Notes: CS = customer satisfaction, IN = innovativeness capability, PQ = product quality, TQ = Tobin's q, and SR = stock return

term of CSR × product quality significantly affects both Tobin's q and stock return, though the interaction term of CSR × innovativeness capability affects only Tobin's q.

To facilitate the interpretation of the moderating effects, Figure 2, Panel A, illustrates the relationship between CSR and Tobin's q for firms with low or high innovativeness capability (see Aiken and West 1991, pp. 12–14). Figure 2, Panel A, suggests that firms with low innovativeness capabilities generate negative market value from CSR, whereas firms with high innovativeness generate positive market value from CSR. However, Figure 2, Panel B, shows that though firms with high product quality generate positive market value from CSR (the upward-sloping line), firms with low product quality seem not to be penalized in terms of generating market value from CSR (the rather flat line). As such, overall, we find support for H_3 when we use innovativeness capability as the measure of corporate abilities, but we find only partial support for H_3 when we use product quality as the measure of corporate abilities.

Results for the Mediating Role of Customer Satisfaction in the Moderated Relationships

H_4 predicted that customer satisfaction would mediate the moderated relationships in H_3. Although a test of this combination of mediation and moderation is somewhat complicated, Baron and Kenny (1986, p. 1179) recommend a practical approach. Essentially, it is similar to the four conditions of mediation we described previously, but it requires entering the interactions items (CSR × innovativeness capability and CSR × product quality) rather than the main effect of CSR. More specifically, to establish mediated moderation, four specific conditions must be met: (1) The interaction variables (CSR × innovativeness capability and CSR × product quality) should significantly influence the mediator (customer satisfaction); (2) the mediator should significantly influence the dependent variable (market value); (3) the interaction variables (CSR × innovativeness capability and CSR × product quality) should significantly influence the dependent variable (market value); and (4) after we control for the mediator variable (customer satisfaction), the impact of the interaction variables (CSR × innovativeness capability and CSR × product quality) on the dependent variable (market value) should be no longer significant (for full mediation) or reduced in strength (for partial mediation) (Baron and Kenny 1986, p. 1179). Following this advice, prior studies in both strategy (Shin and Zhou 2003) and marketing (Andrews et al. 2004; Handelman and Arnold 1999) have tested hypotheses combining mediation and moderation.

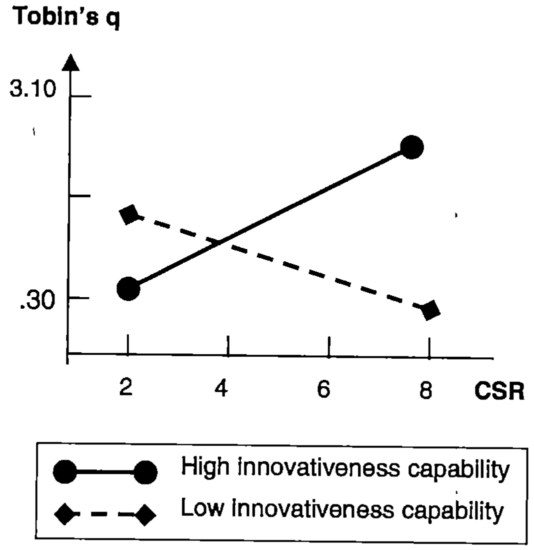
Because the second and third conditions are met, when testing H_1–H_3, we need to check only for the first and fourth conditions. The significant paths from these interaction terms to satisfaction in Model 1 (Table 4) suggest that the first condition is also met. In addition, entering the mediator of customer satisfaction indeed decreases the impact of these interaction terms from Model 2 to Model 3

---

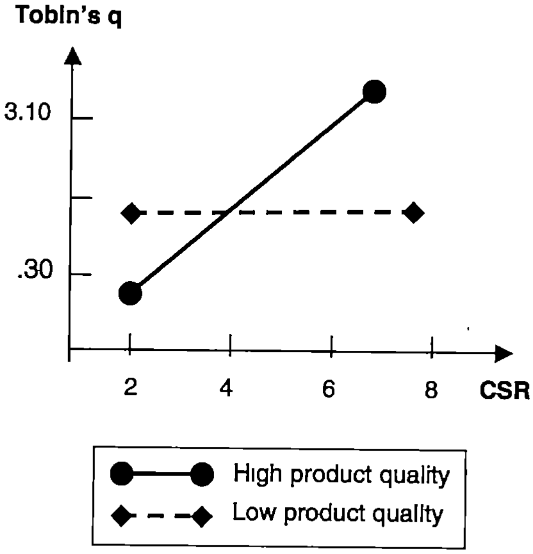
capability, and product quality explained significantly more variance of market value beyond the main effects, adding 6% more variance for Tobin's q and 5% more variance for stock return.

FIGURE 2
The Moderated Effect of CSR on Market Value

A: The Moderating Role of Innovativeness Capacity



B: The Moderating Role of Product Quality



in Table 4. In particular, the impact of CSR × innovativeness capability on Tobin's q is no longer significant, suggesting full mediation (this is not the case for stock return, in which the coefficients in both Model 2 and Model 3 are insignificant). In addition, the impact of CSR × product quality on both Tobin's q and stock return is diminished (but still significant), indicating partial mediation. Thus, these results suggest that the moderation relationships in H_3 are only partially mediated by customer satisfaction, in support of H_4.[8]

Rival Models and Alternative Explanations

We conducted additional analyses and ruled out several competing explanations. Because the aforementioned results suggest partial mediation, we fit several additional SEMs with different partial mediation effects (step-by-step adding/removing of individual paths from the predictive variables of CSR, innovativeness, and product quality to the predicted variables of Tobin's q and stock return). The path estimates of the best-fit partial mediation model appear in Figure 3 and in the last column (Model 4) in Table 4. An examination of Figure 3 suggests three insights. First, although the main effect of CSR is fully mediated by customer satisfaction, CSR has an interaction effect with product quality that is not fully, but rather partially, mediated by customer satisfaction (in other words, this interaction effect between CSR and product quality directly influences Tobin's q and stock return). Second, a firm's product quality and innovativeness both have direct and indirect (through customer satisfaction) influence on Tobin's q performance, which is consistent with prior literature (Dutta, Narasimhan, and Rajiv 1999; Fornell et al. 1996; Rao, Agarwal, and Dahlhoff 2004). Third, we can reject several alternative explanations, including the conjectures that the impact of product quality on market value is fully mediated by satisfaction and that innovativeness capability influences only firm performance but not intermediate outcomes, such as customer satisfaction.

Furthermore, we ruled out several rival models. For example, as we report in Table 4, our SEM results suggest that the partial-mediation SEM (Model 4) fits the data better than the full-mediation SEM (Model 1; $\chi^2_{\text{diff}} = 29.48$, $\text{d.f.}_{\text{diff}} = 5$, $p < .05$) and that the full-mediation SEM fits the data better than the nonmediation SEM (Model 3; $\chi^2_{\text{diff}} = 17.05$, $\text{d.f.}_{\text{diff}} = 4$, $p < .05$). Another criterion for SEM comparison is the number of significant parameters (Morgan and Hunt 1994; Selnes and Sallis 2003). We find that the rival models with full mediation and nonmediation generated fewer significant path estimates. Thus, our hypothesized partial-mediation model fits the data better than competing models in terms of both the relative predictive power of the overall model and the relative number of significant path estimates.

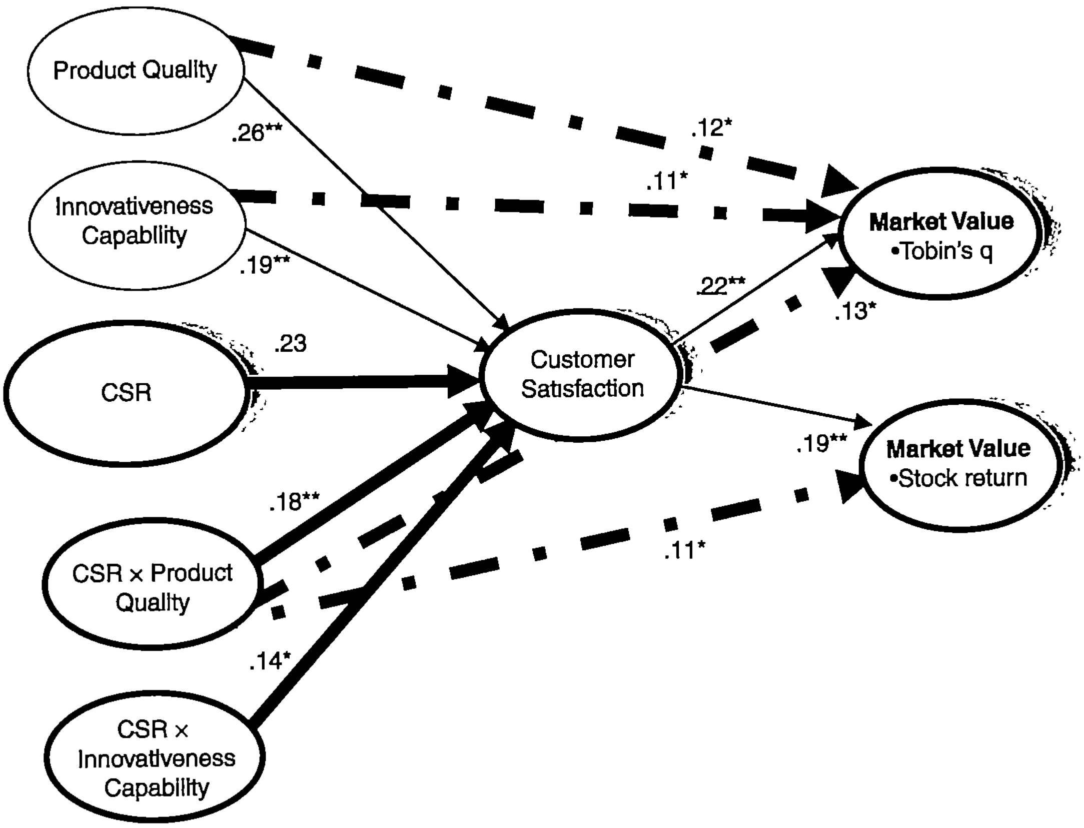
Discussion

How is CSR related to firm market value, and why do CSR initiatives result in financial gains for some firms but losses for others? Our study suggests that the answer to these questions is twofold: (1) CSR affects market value partially through the mediator of customer satisfaction, and (2) returns to CSR can be both positive and negative depending

---

[8] Note that Baron and Kenny (1986, p. 1179) label the relationships we tested as "mediated moderation," which means that controlling the mediator makes the influences of CSR × innovativeness capability and CSR × product quality no longer significant or less significant. This is different from "moderated mediation," in which the moderators should also moderate the mediator–performance linkage (Baron and Kenny 1986, p. 1179). A pictorial illustration of the differences between mediated moderation and moderated mediation can be found in the work of Handelman and Arnold (1999, p. 38).

FIGURE 3
SEM Results of Best-Fit Partial Mediation Model



$^{*}p < .05$, one-tailed test.
$^{**}p < .01$, one-tailed test.
Notes: Bolded paths are hypothesized relationships. Dashed paths are supported partial mediation results. We assessed the significance of all SEM path coefficients through a bootstrapping approach with 1000 resamples.

on the levels of a firm's corporate abilities. Based on a comprehensive secondary data set, our results show that customer satisfaction plays a significant role in the relationship between CSR and firm market value and that a proper combination of both CSR initiatives and product-related abilities is important. These results have implications for both marketing theory and practice.

Before presenting the implications, we note that FAMA's survey-based measure of CSR is an important limitation of this article. As we detailed in the "Data and Variables" section, the FAMA ratings are one possible source of CSR information and thus restrict our analysis and conclusion. To inspire greater confidence in our findings, further research should also attempt to replicate and extend our analysis with alternative measures of CSR. For example, measuring direct spending on CSR initiatives with a large-scale record of CSR monetary expenses across many firms (if obtained reliably from third-party agencies or firms' own reporting; see Margolis and Walsh 2003; Orlitzky, Schmidt, and Rynes 2003; Tsoutsoura 2004) would put CSR on par with measures such as advertising and R&D investments. A clear advantage of this direct approach is that marketing researchers would be able to compare and contrast the financial returns to these different types of spending in an ideal way (i.e., by uncovering the relative, incremental, and synergistic impact of CSR, advertising, and R&D on a firm's market value).

Implications for Marketing Theory

Although CSR has been linked to customer responses (e.g., Bhattacharya and Sen 2004; Brown 1998; Brown and Dacin 1997), this was the first marketing study to explore the relationship between CSR and market value. Our work extends the research stream on the outcomes of CSR from perceived customer responses based on hypothetical lab experiments toward eventual financial returns based on large-scale secondary data. It provides a direct answer to the calls for efforts that link CSR to a firm's stock performance (Berens, Van Riel, and Van Bruggen 2005; Luo and Donthu 2006; Rust, Lemon, and Zeithaml 2004). Indeed, Brown and Dacin (1997, p. 68) note that "we do all good things,... but we don't know if we get anything out of it." The findings pertaining to the significant influence of CSR on a firm's Tobin's q and stock return attest to the financial value of CSR programs as strategic initiatives. Thus, future marketing research should examine a wider spectrum of the benefits of CSR, ranging from perception-based outcomes to archive-based financial returns.

A more important contribution of this research is that we identified a route through which CSR is related to a firm's market value. Our results of the significant CSR → customer satisfaction → market value causal chain suggest that a firm's CSR helps build a satisfied customer base and that customer satisfaction partially mediates the financial returns to CSR. This mediating role of customer satisfaction is important for two reasons. First, it extends the CSR literature by uncovering a previously ignored outcome (i.e., customer satisfaction) of CSR. Although prior work has noted that CSR should affect various kinds of consumer responses, customer satisfaction has not yet been explicitly examined as one such outcome. Second, it also extends the research stream on customer satisfaction (Anderson, Fornell, and Mazvancheryl 2004; Fornell 1992) by uncovering the antecedents (i.e., CSR) of customer satisfaction. Although an emerging research strand has examined the outcomes of customer satisfaction (Anderson, Fornell, and Mazvancheryl 2004; Anderson Fornell, and Rust 1997; Fornell et al. 2006; Mittal et al. 2005), efforts have rarely been undertaken to examine factors that increase or decrease customer satisfaction. Overall, this chained relationship from CSR to customer satisfaction to a firm's market value suggests that achieving customer satisfaction represents one of the underlying pathways through which the financial potential of CSR is realized and capitalized. The notion that the extent to which CSR is beneficial to the firm is determined by how much CSR builds a satisfied customer base points further research in a more precise direction when evaluating the ultimate financial impact of CSR.

Furthermore, our findings suggest that the financial returns to CSR are not the same, but rather are different, across firms with different internal situations. In particular, our finding that the positive financial returns to CSR are amplified in firms with higher product quality indicates that a proper mix or combination of external CSR initiatives and internal corporate abilities likely generates and sustains financial value for the firm. In this sense, we provide empirical evidence for the resource-based view. That is, in support of the resource-based view (Barney 1986; Penrose 1959; Wernerfelt 1984) and marketing capability (Day 1994; Vorhies and Morgan 2005) literature, we find that a firm's sustainable competitive advantages indeed result from a complementary "bundle" of valuable internal (corporate abilities) and external (CSR initiatives) assets. Thus, further research is encouraged to go beyond the simple, universal effects of CSR and explore contextual situations that moderate the associations between CSR and market value.

Finally, existing marketing research has been enthusiastic about the positive benefits of CSR, but unfortunately, it has potential negative outcomes (for an exception, see Sen and Bhattacharya 2001). Our research indicates that CSR can harbor a dark side. That is, in firms that are less innovative in nature, CSR may decrease customer satisfaction levels and ultimately reduce the firm's financial returns. This finding of the negative returns to CSR in the low-innovativeness condition can be understood from the perspective of competitive signaling theory (Caves and Porter 1977; Stigler 1961).

In particular, this theory holds that low (high) innovativeness competency in firms may serve as a cue of inferior (superior) competitiveness to corporate stakeholders, thus signaling weaker (stronger) future performance to financial investors in the marketplace. In the light of signaling theory, we conjecture that though CSR may help firms obtain institutional legitimacy (i.e., by being socially responsive and supportive), firms that are less innovative in meeting customer needs may send a negative signal of incorrect strategic choice and misguided firm priorities in the market that contaminates and degrades this legitimacy (DiMaggio and Powell 1983; Scott 1987). The resulting costs of signaled noncompetitiveness in the market may outweigh the benefits of CSR and thus lead to negative market value. Conceivably, consumers may view CSR activities in firms with low asset specificity as opportunistic (i.e., manipulative and misleading with disguised selling purposes), which causes CSR to backfire and leads to consumer boycotts (Sen and Bhattacharya 2001; Smith 2003). It is also possible that firms that are low in corporate abilities likely invest in less influential and pure cost-adding CSR activities, such as cash donations. In contrast, firms that are high in corporate abilities implement "smarter" CSR strategies that are relatively idiosyncratic and thus generate more long-term financial benefits. We call for further investigation of possible explanations of the observed asymmetric returns to CSR.

Implications for Marketing Practice

Marketers have pondered whether companies should take a more strategic tack on CSR and how "doing good" can contribute to their bottom line (Brown and Dacin 1997; Sen and Bhattacharya 2001). These are important issues that have strong managerial implications because prudent practitioners face tough choices in allocating their limited resources and in prioritizing different strategic initiatives.

> Even evangelists such as Nardelli [chief executive officer of Home Depot] stop short of saying that companies should divert money from other strategic priorities to support [CSR]. But at corporations like Home Depot and [General Electric], good works are being bred into Big Business. 'It's just the right thing to do,' says Nardelli. Good PR? Sure. Money well spent? The goodwill refund could be in the mail. (Grow, Hamm, and Lee 2005, p. 78)

Our finding that CSR contributes positively to market value suggests that managers can obtain competitive advantages and reap more financial benefits by investing in CSR. To be more specific, we calculated that for a typical company in our sample with an average market value of approximately $48 billion, one unit increase of CSR ratings would result in approximately $17 million more profits on average in subsequent years, a substantial increase of financial returns.

Indeed, companies should realize that CSR initiatives can represent a robust public relations strategy, particularly in the current market environment in which stakeholders, such as customers (and employees), may have strong social concerns. Creative executives at Home Depot, IBM, Wal-Mart, General Electric, and Cisco are engaging in "smarter corporate giving" than merely writing checks (Berner 2005, p. 68). For example, Home Depot donated 2 million hours to various types of community services, and IBM gave away more than 100 specialized business applications (i.e., translation servers changing English e-mails into Spanish messages) in heavily Latino-populated schools and community groups. Closer examination of the CSR portfolios of some of the top- and bottom-rated firms in terms of CSR sheds additional light on how managers may derive positive market returns from CSR and/or avoid negative returns. That is, many of the firms at the top of the CSR heap (e.g., United Parcel Service, Alcoa, Verizon Communications) seem to have integrated CSR tightly with their business strategies. For example, these firms invest in a host of employee-related initiatives, such as education and safety, that engender identification and instill pride among employees, all of which influence customer satisfaction and market value. Moreover, these firms have employee volunteerism programs in which employees are visible contributors to the local communities. This helps capture customers' favorable attention.

In contrast, firms at the bottom of the CSR heap, such as Toys 'R' Us and Mitsubishi Motors, seem to be perceived as "irresponsible" by dint of mistreating workers and/or concealing product defect information. Such negative actions tend to receive media coverage in today's scrutiny-intensive world. Viewed in conjunction with our results, these examples suggest that managers should not only "get their house in order" to avoid negative returns to CSR but also adopt an integrated, strategic perspective and allocate resources to CSR programs for superior market performance. After all, "it is no longer an option [for companies] to sit on the sidelines" (Grow, Hamm, and Lee 2005, p. 77; Smith 2003).

Our findings that CSR increases customer satisfaction, which in turn leads to positive financial returns, may improve managers' understanding of why CSR matters. In particular, marketers may have already known that CSR helps promote external social benefits, such as public goodwill outside the firm (Houston and Johnson 2000; McGuire, Sundgren, and Schneeweis 1988), which can polish a firm's reputation in the presence of corporate scandals or regulatory scrutiny. In addition, CSR can boost internal employee morale and commitment within the firm (Godfrey 2005; McGuire, Schneeweis, and Branch 1990) and attract more capable, young talents who are trying to "marry their work and nonwork lives" (Grow, Hamm, and Lee 2005; for detailed benefits of CSR and cause-related marketing, see Varadarajan and Menon 1988, p. 60). Importantly, we suggest an additional insight to managers: CSR initiatives also influence customers' satisfaction levels, which ultimately lead to higher market returns. To managers, this means that building satisfaction is an important intermediate step in converting CSR into financial gains.

However, our findings of the boundary conditions of the returns to CSR suggest that managers should not ignore the inherent traps and pitfalls of CSR. For example, we show that firms are not always able to benefit from CSR actions. When companies are not innovative, our findings indicate that CSR actually decreases their market return. Thus, CSR seems to be a double-edged sword; without proper support of corporate abilities, such as innovativeness, CSR can be harmful to firm performance. Indeed, when "doing better at doing good" (Bhattacharya and Sen 2004, p. 9), it is important for managers to consider CSR initiatives in the light of the firm's corporate abilities. In particular, less innovative companies may be better off financially by avoiding CSR actions. Managers should understand that a misalignment of CSR with internal factors can be detrimental and lead to decreased market value. As a consequence, marketers need to examine carefully the organizational context in totality before implementing CSR initiatives.

In conclusion, this research suggests a more nuanced understanding of the market returns to CSR initiatives. Our findings seem to indicate that "doing good" has complicated implications and that customer satisfaction plays an important mediating role in the relationship between CSR and firm market value.

REFERENCES

Aaker, David and Robert Jacobson (1994), "The Financial Information Content of Perceived Quality," *Journal of Marketing Research*, 31 (May), 191–201.

——— and ——— (2001), "The Value Relevance of Brand Attitude in High-Technology Markets," *Journal of Marketing Research*, 38 (November), 485–93.

Ahearne, Michael, C.B. Bhattacharya, and Thomas Gruen (2005), "Antecedents and Consequences of Customer-Company Identification: Expanding the Role of Relationship Marketing," *Journal of Applied Psychology*, 90 (3), 574–85.

Aiken, Leona and Stephen G. West (1991), *Multiple Regression: Testing and Interpreting Interactions*. Thousand Oaks, CA: Sage Publications.

Anderson, Eugene W., Claes Fornell, and Sanal K. Mazvancheryl (2004), "Customer Satisfaction and Shareholder Value," *Journal of Marketing*, 68 (October), 172–85.

———, ———, and Roland T. Rust (1997), "Customer Satisfaction, Productivity, and Profitability: Differences Between Goods and Services," *Marketing Science*, 16 (2), 129–45.

Anderson, James C. and David W. Gerbing (1988), "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin*, 103 (3), 411–23.

Andrews, J. Craig, Richard Netemeyer, Scot Burton, Paul Moberg, and Ann Christainsen (2004), "Understanding Adolescent Intentions to Smoke: An Examination of Relationships Among Social Influences, Prior Trial Behaviors, and Antitobacco Campaign Advertising," *Journal of Marketing*, 68 (July), 110–23.

Aupperle, K., A. Carroll, and J. Hatfield, (1985), "An Empirical Examination of the Relationship Between Corporate Social Responsibility and Profitability," *Academy of Management Journal*, 28 (2), 446–63.

Barney, Jay B. (1986), "Strategic Factor Markets: Expectations, Luck, and Business Strategy," *Management Science*, 32 (10), 1231–41.

Baron, Reuben M. and David A. Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (6), 1173–82.

Bhattacharya, C.B., Hayagreeva Rao, and Mary Ann Glynn (1995), "Understanding the Bond of Identification: An Investigation of Its Correlates Among Art Museum Members," *Journal of Marketing*, 59 (October), 46–57.

——— and Sankar Sen (2003), "Consumer–Company Identification. A Framework for Understanding Consumers' Relationships with Companies," *Journal of Marketing*, 67 (April), 76–88.

——— and ——— (2004), "Doing Better at Doing Good," *California Management Review*, 47 (1), 9–24.

Berens, Guido, Cees B.M. van Riel, and Gerrit H. van Bruggen (2005), "Corporate Associations and Consumer Product Responses: The Moderating Role of Corporate Brand Dominance," *Journal of Marketing*, 69 (July), 35–18.

Berner, Robert (2005), "Smarter Corporate Giving," *BusinessWeek*, (November 28), 68–76.

Bluedorn, Allen (1982), "A Unified Model of Turnover from Organizations," *Human Relations*, 35 (February), 135–53.

Boles, James S., Mark W. Johnston, and Joseph F Hair Jr. (1997), "Role Stress, Work-Family Conflict and Emotional Exhaustion: Inter-Relationships and Effects on Some Work-Related Consequences," *Journal of Personal Selling & Sales Management*, 17 (1), 17–28.

Bolton, Ruth N. and James H. Drew (1991), "A Longitudinal Analysis of the Impact of Service Changes on Customer Attitudes," *Journal of Marketing*, 55 (January), 1–9.

Brown, Tom J. (1998), "Corporate Associations in Marketing: Antecedents and Consequences," *Corporate Reputation Review*, 1 (3), 215–33.

——— and Peter A. Dacin (1997), "The Company and the Product: Corporate Associations and Consumer Product Responses," *Journal of Marketing*, 61 (January), 68–84.

Buzzell, Robert D., Bradley T. Gale, and Ralph G.M. Sultan (1975), "Market Share: A Key to Profitability," *Harvard Business Review*, 53 (1), 97–106.

Caves, R.E. and Michael E. Porter (1977), "From Entry Barriers to Mobility Barriers," *Quarterly Journal of Economics*, 91 (2), 421–34.

Chauvin, Keith W. and Mark Hirschey (1993), "Advertising, R&D Expenditures, and the Market Value of the Firm," *Financial Management*, 22 (Winter), 128–40.

Christmann, P. (2000), "Effects of 'Best Practices' of Environmental Management on Cost Advantage: The Role of Complementary Assets," *Academy of Management Journal*, 43 (4), 663–80.

Cho, Hee-Jae and Vladimir Pucik (2005), "Relationship Between Innovativeness, Quality, Growth, Profitability, and Market Value," *Strategic Management Journal*, 26 (6), 555–75.

Daub, Claus-Heinrich and Rudolf Ergenzinger (2005), "Enabling Sustainable Management Through a New Multi-Disciplinary Concept of Customer Satisfaction," *European Journal of Marketing*, 39 (9–10), 998–1012.

Day, George (1994), "The Capabilities of Market-Driven Organizations," *Journal of Marketing*, 58 (October), 37–52.

DiMaggio, P.J. and W.W. Powell (1983), "The Iron Cage Revisited: Instructional Isomorphism and Collective Rationality in Organizational Fields," *American Sociological Review*, 48 (2), 17–60.

Drucker, Peter F. (1993), *Post-Capitalist Society*. New York: HarperCollins.

Drumwright, Minette E. (1996), "Company Advertising with a Social Dimension: The Role of Noneconomic Criteria," *Journal of Marketing*, 60 (October), 71–87.

Dutta, Shantanu, Om Narasimhan, and Surendra Rajiv (1999), "Success in High-Technology Markets: Is Marketing Capability Critical?" *Marketing Science*, 18 (4), 547–68.

Erickson, Gary and Robert Jacobson (1992), "Gaining Comparative Advantage Through Discretionary Expenditures: The Returns to R&D and Advertising," *Management Science*, 38 (September), 1264–79.

Fombrun, Charles and Mark Shanley (1990), "What's in a Name? Reputation Building and Corporate Strategy," *Academy of Management Journal*, 33 (2), 233–58.

Fornell, Claes (1992), "A National Customer Satisfaction Barometer: The Swedish Experience," *Journal of Marketing*, 6 (January), 1–21.

———, Michael D. Johnson, Eugene W. Anderson, Jaesung Cha, and Barbara Bryant (1996), "The American Customer Satisfaction Index: Description, Findings, and Implications," *Journal of Marketing*, 60 (October), 7–18.

——— and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 19 (February), 39–50.

———, Sunil Mithas, Forrest V. Morgeson III, and M.S. Krishnan (2006), "Customer Satisfaction and Stock Prices: High Returns, Low Risk," *Journal of Marketing*, 70, (January), 3–14.

Fortune (2005), "America's Most Admired Companies," (March 7), 68

Gatignon, Hubert and Jean-Marc Xuereb (1997), "Strategic Orientation of the Firm and New Product Performance," *Journal of Marketing Research*, 34 (February), 77–90.

Godfrey, Paul C (2005), "The Relationship Between Corporate Philanthropy and Shareholder Wealth: A Risk Management Perspective," *Academy of Management Review*, 30 (4), 777–98.

Griffin, Abbie and J.R. Hauser (1996), "Integrating R&D and Marketing: A Review and Analysis of the Literature," *Journal of Product Innovation Management*, 13 (3), 191–215.

Grow, Brian, Steve Hamm, and Louise Lee (2005), "The Debate Over Doing Good," *BusinessWeek*, (August 15), 76–78.

Gruca, Thomas S. and Lopo L. Rego (2005), "Customer Satisfaction, Cash Flow, and Shareholder Value," *Journal of Marketing*, 69 (July), 115–30.

Gürhan-Canli, Zeynep and Rajeev Batra (2004), "When Corporate Image Affects Product Evaluations: The Moderating Role of Perceived Risk," *Journal of Marketing Research*, 41 (May), 197–205.

Handelman, Jay M. and Stephen J. Arnold (1999), "The Role of Marketing Actions with a Social Dimension: Appeals to the Institutional Environment," *Journal of Marketing*, 63 (July), 33–48.

Holmbeck, Grayson (1997), "Toward Terminological, Conceptual, and Statistical Clarity in the Study of Mediators and Moderators: Examples from the Child-Clinical and Pediatric Psychology Literatures," *Journal of Consulting and Clinical Psychology*, 65 (4), 599–610.

Homburg, Christian, Nicole Koschate, and Wayne D. Hoyer (2005), "Do Satisfied Customers Really Pay More? A Study of the Relationship Between Customer Satisfaction and Willingness to Pay," *Journal of Marketing*, 69 (April), 84–97.

Houston, Mark and Shane Johnson (2000), "Buyer–Supplier Contracts Versus Joint Ventures: Determinants and Consequences of Transaction Structure," *Journal of Marketing Research*, 37 (February), 1–15.

Jayachandran, Satish, Subhash. Sharma, Peter Kaufman, and Pushkala Raman (2005), "The Role of Relational Information Processes and Technology Use in Customer Relationship Management," *Journal of Marketing*, 69 (October), 177–92.

Johnston, Mark, Charles Futrell, A. Parasuraman, and William Black (1990), "A Longitudinal Assessment of the Impact of Selected Organizational Influences on Salespeople's Organizational Commitment During Early Employment," *Journal of Marketing Research*, 27 (August), 333–44.

Joseph, Kissan and Vernon J. Richardson (2002), "Free Cash Flow, Agency Costs, and the Affordability Method of Advertising Budgeting," *Journal of Marketing*, 66 (January), 94–107.

Kim, W.C. and R. Mauborgne (1997), "Value Innovation: The Strategic Logic of High Growth," *Harvard Business Review*, 75 (1), 102–12.

Kleinschmidt E.J. and R.G. Cooper (1991), "The Impact of Product Innovativeness on Performance," *Journal of Product Innovation Management*, 8 (4), 240–51.

Kotler, Philip and Nancy Lee (2004), *Corporate Social Responsibility: Doing the Most Good for Your Company and Your Cause*. New York: John Wiley & Sons.

Lee, Ruby and Rajdeep Grewal (2004), "Strategic Responses to New Technologies and Their Impact on Firm Performance," *Journal of Marketing*, 68 (October), 157–71.

Li, Haiyang and Kwaku Atuahene-Gima (2001), "Product Innovation Strategy and the Performance of New Technology Ventures in China," *Academy of Management Journal*, 44 (6), 1123–34.

Lichtenstein, Donald R., Minette E. Drumwright, and Bridgette M. Braig (2004), "The Effect of Corporate Social Responsibility on Customer Donations to Corporate-Supported Nonprofits," *Journal of Marketing*, 68 (October), 16–32.

Luo, Xueming and Naveen Donthu (2006), "Marketing's Credibility: A Longitudinal Investigation of Marketing Communication Productivity and Shareholder Value," *Journal of Marketing*, 70 (October), 70–91.

Maignan, Isabelle, O.C. Ferrell, and Linda Ferrell (2005), "A Stakeholder Model for Implementing Social Responsibility in Marketing," *European Journal of Marketing*, 39 (9–10), 956–77.

March, J.G. (1991), "Exploration and Exploitation in Organizational Learning," *Organization Science*, 2 (1), 71–87.

Margolis, Joshua D. and James P. Walsh (2003), "Misery Loves Companies: Rethinking Social Initiatives by Business," *Administrative Science Quarterly*, 38 (3), 268–305.

McGuire, Jean, Thomas Schneeweis, and Ben Branch (1990), "Perceptions of Firm Quality: A Cause or Result of Firm Performance," *Journal of Management*, 16 (1), 167–80.

———, Alison Sundgren, and Thomas Schneeweis (1988), "Corporate Social Responsibility and Firm Financial Performance," *Academy of Management Journal*, 31 (4), 854–72.

Mithas, S , M.S. Krishnan, and Claes Fornell (2005a), "Effect of Information Technology Investments on Customer Satisfaction: Theory and Evidence," working paper, Ross School of Business, University of Michigan.

———, ———, and ——— (2005b), "Why Do Customer Relationship Management Applications Affect Customer Satisfaction?" *Journal of Marketing*, 69 (October), 201–209.

Mittal, Vikas, Akin Sayrak, Pandu Tadikamalla, and Eugene Anderson (2005), "Dual Emphasis and the Long-Term Financial Impact of Customer Satisfaction," *Marketing Science*, 24 (4), 544–55.

Mizik, Natalie and Robert Jacobson (2003), "Trading Off Between Value Creation and Value Appropriation: The Financial Implications of Shifts in Strategic Emphasis," *Journal of Marketing*, 67 (January), 63–76.

Morgan, Neil A. and Lopo Leotte do Rego (2006), "The Value of Different Customer and Loyalty Metrics in Predicting Business Performance," *Marketing Science*, forthcoming.

Morgan, Robert M. and Shelby D. Hunt (1994), "The Commitment–Trust Theory of Relationship Marketing," *Journal of Marketing*, 58 (July), 20–38.

Murthi, B.P.S., Kannan Srinivasan, and Gurumurthy Kalyanaram (1996), "Controlling for Observed and Unobserved Managerial Skills in Determining First-Mover Market Share Advantages," *Journal of Marketing Research*, 33 (August), 329–37.

Netemeyer, Richard G., James G. Maxham III, and Chris Pullig (2005), "Conflicts in the Work–Family Interface: Links to Job Stress, Customer Service Employee Performance, and Customer Purchase Intent," *Journal of Marketing*, 69 (April), 130–43.

Nunnally, Jum C. (1978), *Psychometric Theory*, 2d ed. New York: McGraw-Hill.

Oliver, Richard L. (1980), "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research*, 17 (November), 460–69.

Orlitzky, Marc, Frank Schmidt, and Sara L. Rynes (2003), "Corporate Social and Financial Performance: A Meta-Analysis," *Organization Studies*, 24 (3), 403–441.

Pan, Xing, Brian Ratchford, and Venkatesh Shankar (2002), "Can Price Dispersion in Online Markets Be Explained by Differences in E-tailer Service Quality?" *Journal of the Academy of Marketing Science*, 30 (4), 433–45.

Pava, Moses L. and Joshua Krausz (1996), *Corporate Social Responsibility and Financial Performance: The Paradox of Social Cost*. Westport, CT: Quorum Books.

Penrose, Edith T. (1959), *The Theory of the Growth of the Firm*. Oxford: Basil Blackwell.

Rao, Vithala, Manoj K. Agarwal, and Denise Dahlhoff (2004), "How Is Manifest Branding Strategy Related to the Intangible Value of a Corporation?" *Journal of Marketing*, 68 (October), 126–41.

Roberts, Peter and Grahame Dowling (2002), "Corporate Reputation and Sustained Superior Financial Performance," *Strategic Management Journal*, 23 (2), 1077–1093.

Rust, Roland, Katherine Lemon, and Valarie A. Zeithaml (2004), "Return on Marketing: Using Customer Equity to Focus Marketing Strategy," *Journal of Marketing*, 68 (January), 109–124.

———, Christine Moorman, and Peter R. Dickson (2002), "Getting Return on Quality: Cost Reduction, Revenue Expansion, or Both?" *Journal of Marketing*, 66 (October), 7–24.

Schumpeter, J.A. (1934), *The Theory of Economic Development*. Cambridge, MA: Harvard University Press.

Scott, W. Richard (1987), "The Adolescence of Institutional Theory," *Administrative Science Quarterly*, 32 (December), 493–511.

Selnes, Fred and James Sallis (2003), "Promoting Relationship Learning," *Journal of Marketing*, 67 (July), 80–95.

Sen, Sankar and C.B. Bhattacharya (2001), "Does Doing Good Always Lead to Doing Better? Consumer Reactions to Corporate Social Responsibility," *Journal of Marketing Research*, 38 (May), 225–44.

Shin, Shung Jae and Jing Zhou (2003), "Transformational Leadership, Conservation, and Creativity: Evidence from Korea," *Academy of Management Journal*, 46 (6), 703–714.

Smith, Craig N. (2003), "Corporate Social Responsibility: Whether or How?" *California Management Review*, 45 (4), 52–76.

Soloman, R. and K. Hansen (1985), *It's Good Business*. New York: Atheneum.

Srivastava, Rajendra, Tasadduq Shervani, and Liam Fahey (1998), "Market-Based Assets and Shareholder Value: A Framework for Analysis," *Journal of Marketing*, 62 (January), 2–18.

Stigler, G.J. (1961), "The Economics of Information," *Journal of Political Economy*, 69 (3), 104–122.

Szymanski, David M. and David Henard (2001), "Customer Satisfaction: A Meta-Analysis of the Empirical Evidence," *Journal of the Academy of Marketing Science*, 29 (Winter), 16–35.

Tsoutsoura, Margarita (2004), "Corporate Social Responsibility and Financial Performance," working paper, Haas School of Business, University of California, Berkeley.

Varadarajan, P. Rajan and Anil Menon (1988), "Cause-Related Marketing: A Coalignment of Marketing Strategy and Corporate Philanthropy," *Journal of Marketing*, 52 (July) 58–74.

Vargo, Stephen L. and Robert F. Lusch (2004), "Evolving to a New Dominant Logic for Marketing," *Journal of Marketing*, 68 (January), 1–17.

Vorhies, Douglas W. and Neil A. Morgan (2005), "Benchmarking Marketing Capabilities for Sustainable Competitive Advantage," *Journal of Marketing*, 69 (January), 80–94.

Wernerfelt, Birger (1984), "A Resource-Based View of the Firm," *Strategic Management Journal*, 5 (2), 171–80.

Zeithaml, Valarie A. (2000), "Service Quality, Profitability, and the Economic Worth of Customers: What We Know and What We Need to Learn," *Journal of the Academy of Marketing Science*, 28 (1), 67–85.

———, A. Parasuraman, and Leonard L. Berry (1990), *Delivering Quality Service: Balancing Customer Perceptions and Expectations*. New York: The Free Press.

Martin Holzwarth, Chris Janiszewski, & Marcus M. Neumann

The Influence of Avatars on Online Consumer Shopping Behavior

An impediment to Web-based retail sales is the impersonal nature of Web-based shopping. A solution to this problem is to use an avatar to deliver product information. An avatar is a graphic representation that can be animated by means of computer technology. Study 1 shows that using an avatar sales agent leads to more satisfaction with the retailer, a more positive attitude toward the product, and a greater purchase intention. Study 2 shows that an attractive avatar is a more effective sales agent at moderate levels of product involvement, but an expert avatar is a more effective sales agent at high levels of product involvement.

The Internet is a major communication and transaction medium. It is estimated that between 700 million and 950 million people used the Internet in 2004 and that close to 500 million of these people were frequent users (i.e., more than once a week). Approximately 200 million people engaged in retail commerce on the Internet, with worldwide retail sales approaching $70 billion in 2004 (e.g., United States: $44 billion; Europe: $10 billion; Asia: $10 billion). Yearly Internet retail sales growth averaged 30% per year from 2001 to 2005 and is expected to grow at a similar rate in the near future.

Despite the strong growth of Internet sales, this distribution channel has not been as successful as was once projected. In 2000, it was anticipated that as much as 7.8% of U.S. retail sales would be transacted on the Internet in 2005 (Dykema 2000). However, only 2.0% of U.S. retail sales were transacted on the Internet in the fourth quarter of 2004 (U.S. Census Bureau 2005), and an even smaller percentage was transacted in Europe and Asia. The greatest impediment to Internet channel growth has been the low shopper conversion rate, that is, the percentage of visitors to a retail site who actually make a purchase. The conversion rate of Internet shoppers averages only 4.9% among the *top* 100 Internet retailers (Nielsen//NetRatings 2005), a rate significantly lower than experienced by comparable firms using traditional retailing channels. Moreover, research shows that between 65% and 75% of consumers who initiate an online transaction fail to complete the transaction (Mummert & Partner 2001).

Survey research suggests that the most significant inhibitors of online shopping are the absence of pleasurable experiences, social interaction, and personal consultation by a company representative (Barlow, Siddiqui, and Mannion 2004; G&J Electronic Media Services 2001). Consumers report that online companies are impersonal, they feel helpless when shopping virtually in unfamiliar or complex product categories, and they want the customer assistance often found in a conventional shopping environment combined with the convenience of Internet shopping. Thus, improving the Internet shopping experience should improve the conversion rate of potential buyers (Childers et al. 2001).

One approach to increasing the entertainment value, information value, and customer satisfaction of Web-based shopping experiences is to use "avatars" (Barlow, Siddiqui, and Mannion 2004; Redmond 2002). Avatars are virtual characters that can be used as company representatives. Avatars can serve as identification figures, as personal shopping assistants, as Web site guides, or as conversation partners. In these roles, avatars have the potential to fulfill the consumer's desire for a more interpersonal shopping experience. Thus, an electronic shopping agent may be able to ease a consumer's navigation of a Web site or provide personalized information, but an avatar can anthropomorphize the interaction and make the shopping experience more interpersonal. Consequently, the information provided on the Web site should be perceived as more credible, the shopping experience should become more enjoyable, and the likelihood of a purchase should increase.

The goal of this research is to investigate the benefits of using avatars as company representatives on commercial Web sites. Two studies show that avatars positively affect an online shopping experience. In Study 1, an avatar communicator creates a more positive perception of the entertainment value and informativeness of a Web site. Consequently, shoppers are more satisfied with the retailer, more positive about the product, and more likely to purchase the product. These avatar advantages persist even when information content is held constant between an avatar and a no-avatar format. Study 2 finds that moderately involved shoppers are more persuaded by attractive-looking avatars and that highly involved shoppers are more persuaded by expert avatars. Attractive avatars are persuasive because of their

---

Martin Holzwarth is a consultant at Marketing Partner, MP Management Consultants GmbH, Wiesbaden, Germany (e-mail: mh@martinholzwarth.com). Chris Janiszewski is Jack Faricy Professor of Marketing, Warrington College of Business Administration, University of Florida (e-mail: chris.janiszewski@cba.ufl.edu). Marcus M. Neumann is a doctoral student, Department of Business Administration and Marketing II, University of Mannheim (e-mail: marcus.neumann@bwl.uni-mannheim.de). All three authors contributed equally to this work. The authors thank Alan Sawyer for his helpful comments.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

likeability, whereas expert avatars are persuasive because of their credibility.

Avatars

The word "avatar" has its derivation in the ancient Indian language Sanskrit and refers to the embodiment of a deity on earth. Consistent with this original definition, present-day definitions of an avatar refer to a representation of an entity. For example, an avatar is described as "a pictorial representation of a human in a chat environment" (Bahorsky, Graber, and Mason 1998, p. 8) or as "a representation of the user as an animated character in virtual worlds" (Loos 2003, p. 17). There are also technology-oriented approaches to defining avatars. For example, the computer trade press describes avatars as "graphic personifications of computers or processes that run on computers" (Halfhill 1996, p. 69). For our purposes, avatars are defined as general graphic representations that are personified by means of computer technology.

Computer-Mediated Interaction

According to the theory of social response, people tend to react to computer technology as though it is a social entity (Moon 2000, 2003; Reeves and Nass 1996). Whenever computer technology exhibits humanlike behaviors, such as language production, taking turns in conversation, and reciprocal responding, the user is more apt to personify the technology (Moon 2000; Nass et al. 1995a). This tendency to treat a computer as a social entity occurs whether the representation of the computer is the screen, a voice, or an agent (Moon 2000). Thus, when people are confronted with a computer or software program, they have a tendency to engage in a "social response to (these) communication technologies," or what Morkes, Kernal, and Nass (1999) call the SRCT approach.

An early example of a social response to computer technology is provided by Weizenbaum (1976), who exposed experimental participants to a computer-based speech analysis program called ELIZA. This program reformulated inputs from the participants into psychoanalytical questions. The interaction caused participants to develop an emotional relationship to and attribute human characteristics to the computer, even though the participants were fully aware that the computer did not possess a personality (Morkes, Kernal, and Nass 1999). More recent evidence shows that users respond to computers as they do to people in terms of psychosocial phenomena, such as personality, politeness, and flattery (Moon 2003; Nass et al. 1995a; Nass et al. 1995b; Nass and Steur 1993).

Avatar-Mediated Interaction

Research on computer-mediated interactions persuasively demonstrates that computer technology can be personified. Yet the research is silent with respect to the benefits of using an avatar to represent the technology. However, the research on mediated communication may provide insight into this issue. This stream of research suggests that mediated communications (e.g., radio, television) are less persuasive than interpersonal communication because mediated communications lack reciprocity (Horton and Wohl 1956; Rubin, Perse, and Powell 1985). There is also evidence that increasing the reciprocity of a technology-based interaction through verbal disclosures increases compliance with requests made by the technology (Moon 2000). We posit that simply having an avatar pictured during the human–computer interaction will make the interaction feel more conversational and reciprocal. Thus, adding an avatar to a decision support system on a retail Web site will increase the effectiveness of the Web site. This prediction is consistent with recent evidence that relationships are more likely to develop if the computer technology is represented using human forms (Trogemann 2003).

Avatars as Persuasion Agents

If avatars enhance the personification of a technology, they should influence the purchase process in a manner similar to human sales agents. Human sales agents have been shown to increase satisfaction with a retailer, enhance attitudes toward products sold by the retailer, and increase the consumer's intention to buy (Katz and Lazarsfeld 1955; Webster 1968). Thus, avatars should have a similar impact on the shopping experience (Redmond 2002).

> H_1: Avatar-mediated communication has a positive effect on (a) satisfaction with the retailer, (b) attitude toward the product, and (c) purchase intention.

It is also important to understand how avatar-mediated communications can influence purchase behavior. First, there is a long history of evidence that personal communication is more effective than mass-media communication (Williams 1977). Second, there is a long history of evidence that personal communication is more complex, adaptive, and satisfying than other forms of communication (Allen et al. 2002). Both research streams help explain the effectiveness of face-to-face interactions with sales agents. In general, face-to-face interactions with sales agents are effective not only because sales agents can provide information, be empathic, and build rapport (Barlow, Siddiqui, and Mannion 2004) but also because the agent's information is perceived as more accurate and is more likely to be believed (Soldow and Thomas 1984). In effect, sales agents enhance the value of the information provided and increase the pleasure of the shopping experience, especially in a retailing environment (Reynolds and Beatty 1999). Given that people visit Web sites for entertainment needs and information needs (Eighmey and McCord 1998; Korgaonkar and Wolin 1999), avatar-mediated communications may have an impact similar to human–sales-agent-mediated communications.

> H_2: Avatar-mediated communication has a positive effect on (a) the perceived entertainment value of a Web site and (b) the perceived information value of a Web site.

Finally, we expect that the perceived entertainment and information value of a Web site will mediate the consumer's satisfaction with the retailer, attitude toward the product, and purchase intention. Our expectation of the influence of these mediating factors is related to our previous discussion of deterrents to Internet shopping. Recall that an important deterrent to consumers' use of the Internet for product purchases is the limited range of personalized services available within this medium. The most notable omission is the lack of face-to-face communication (Phillips et al. 1997). Swaminathan and colleagues (1999) observe that many people have a strong desire for social interaction in their shopping experience. If a marketer can positively affect a customer's mood by providing entertaining content, this should have a positive effect on the perception of the company and its products (Brown and Stayman 1992; MacKenzie and Lutz 1989; MacKenzie, Lutz, and Belch 1986). Similarly, if a marketer can make information seem more relevant and important, this should encourage a more positive response toward the retailer and its offerings (Weitz, Sujan, and Sujan 1986).

> H_3: (a) The perceived entertainment value of a Web site and (b) the perceived information value of a Web site mediate the positive influence of an avatar-mediated communication on satisfaction with the retailer, attitude toward the product, and purchase intention.

We summarize these relationships in Figure 1 under the heading "Influence of Avatars."

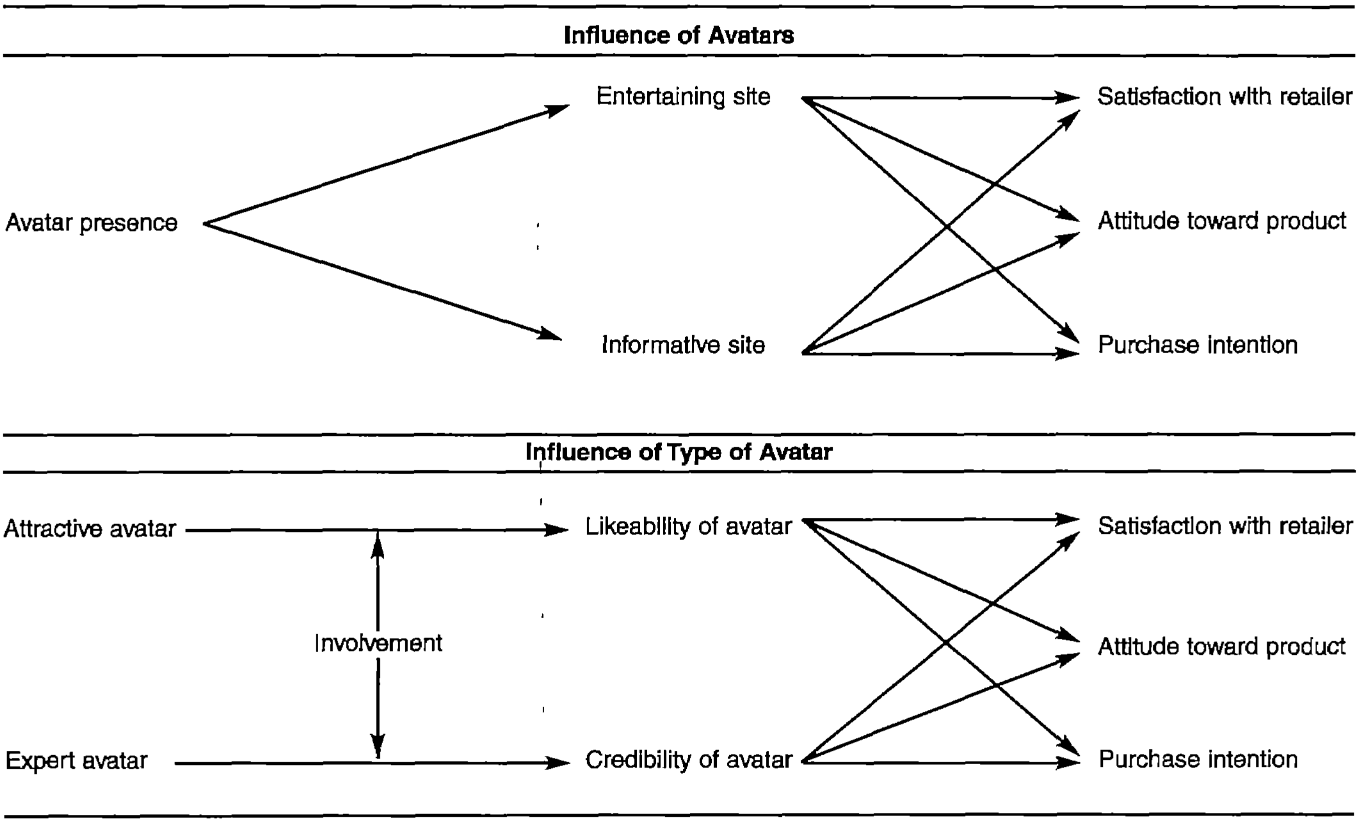
Designing Effective Avatars

Kelman (1961) explains how three processes of social influence affect the message recipient's acceptance of a message. The first process focuses on the recipient's identification with the communicator. In this case, the recipient adopts the attitude of the communicator through imitation or internalization of the message. The *interpersonal attraction* of the communicator is a prerequisite for motivating the recipient to adopt the communicator's position. The second process concerns the recipient's perceptions of the communicator's credibility as a consequence of the *expertise* of the communicator and the trust that develops between the communicator and the recipient. In this case, the credibility of the communicator plays a key role in persuasion. The third process pertains to the compliance of the recipient. The recipient is more compliant if he or she feels controlled by the communicator and/or the communicator can provide rewards and punishments. We focus on the first two processes.

Moderator. We anticipate that the attractiveness and the expertise of an avatar will have an influence on its persuasiveness and that the consumer's involvement with the purchase will moderate the influence of these characteristics. The expectation that the persuasiveness of different types of avatars will vary with involvement is consistent with a long history of research on persuasion by human communicators (see Cooper and Croyle 1984; Wood 2000). For example, the attractiveness of a communicator has been shown to have a greater impact on persuasion at lower levels of involvement (Petty, Cacioppo, and Goldmann 1981), whereas the expertise of a communicator's message has been shown to have a greater impact on persuasion at higher levels of involvement (Petty, Cacioppo, and Heesacker 1981). The positive characteristics of the communicator can generalize to the products being promoted by these people (Caballero and Pride 1984).

The varied influence of communicator attractiveness and communicator expertise may generalize to avatar com-

FIGURE 1
Predicted Relationships



munications. At the lowest level of purchase involvement, there will be little difference between an attractive and an expert avatar because the communication agent and message are irrelevant to the consumer. An avatar's attractiveness will contribute to an avatar's persuasiveness as involvement increases from a low to a moderate level, but it will have less influence on persuasiveness as involvement increases from a moderate to a high level. An avatar's expertise, as exemplified in the message quality, will contribute to an avatar's persuasiveness as involvement increases from a low to a moderate to a high level. In summary, the influence of avatar attractiveness on persuasion is a log function of increasing involvement, whereas the influence of avatar expertise on persuasion is a linear function of increasing involvement (see Figure 2). We expect that these different patterns of response to avatar attractiveness and expertise will allow an attractive avatar to be more persuasive at moderate levels of involvement and an expert avatar to be more persuasive at high levels of involvement. Formally,

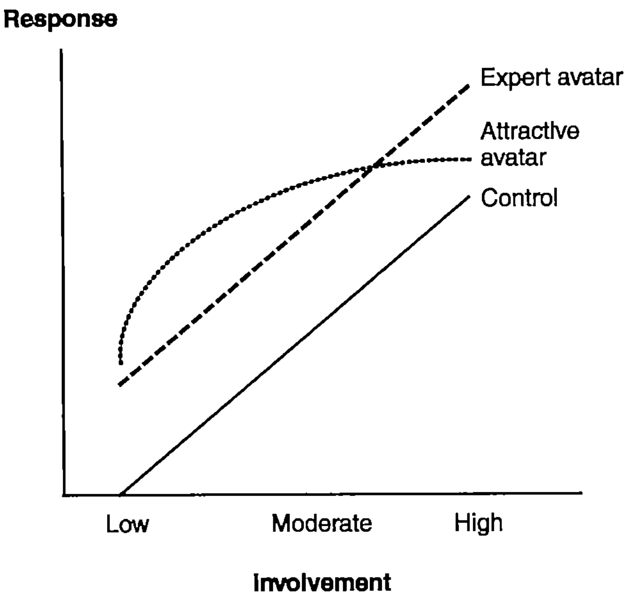
> H_4: The influence of avatar attractiveness on persuasion is a log function of increasing involvement, whereas the influence of avatar expertise on persuasion is a linear function of increasing involvement.

> H_{4a}: An attractive avatar will be more persuasive than an expert avatar at moderate levels of involvement.

> H_{4b}: An expert avatar will be more persuasive than an attractive avatar at high levels of involvement.

Mediators. Whereas H_3 discusses factors (e.g., entertainment value, information value) that might mediate the positive influence of having an avatar present on a Web site, H_4 hypothesizes that different types of avatars (e.g., attractive, expert) may be effective in different situations. It is also possible to speculate about processes that mediate the relative effectiveness of an attractive or expert avatar. For example, interactions with physically attractive people are pleasant and rewarding (Bull and Rumsey 1988). The pleasant feeling that results from an interaction with an attractive person, in turn, promotes the adoption of the communicator's behaviors and attitudes (Chaiken 1979; Dion, Berscheid, and Walster 1972; Joseph 1982; McGuire 1985). In addition, physically attractive people are perceived more favorably on traits typically associated with selling effectiveness (Reingen and Kernan 1993). Thus, the attractiveness of an avatar should influence the likeability of an avatar, and likeability should mediate the degree of persuasion.

FIGURE 2
Predictions



> H_{5a}: The likeability of the avatar mediates the persuasiveness of an attractive avatar relative to an unattractive avatar.

Similarly, the expertise of an avatar should influence perceptions of the credibility of the avatar, and credibility should mediate the degree of persuasion. For our purposes, it is important to note that particularly experienced or knowledgeable people are perceived as experts (Friedman and Friedman 1979; Stäudel 1987) and that expert communicators are perceived as more credible than nonexpert communicators (Brehm, Kassin, and Fein 2005; Hovland, Janis, and Kelley 1953; McGuire 1985). In general, research investigating source expertise in persuasive communications indicates that the source's perceived credibility has a positive impact on persuasion (Horai, Naccari, and Fatoullah 1974; Maddux and Rogers 1980; Mills and Harvey 1972). For example, the expertise of a corporate communication agent affects the perceived credibility of the communication (Freiden 1984; Rubin, Mager, and Friedman 1982) and the associated company (Stephens and Faranda 1993), as well as the acceptance of the message. To the extent that avatars persuade as human communicators do, we hypothesize similar mediation effects for avatars.

> H_{5b}: The perceived credibility of the avatar mediates the persuasiveness of an expert avatar relative to a nonexpert avatar.

We summarize these relationships in Figure 1 under the heading "Influence of Type of Avatar."

Study 1

Study 1 explored the influence of avatars on consumer responses to Web-based merchandising. The purchase scenario was an opportunity to purchase a leisure shoe that could be customized by means of an online consultation. This online consultation was performed using a series of diagnostic screens that solicited the consumer to make product design decisions. The key manipulation was whether the consumer was led to believe that an impersonal software program or an avatar was soliciting information and design recommendations.

Design and Stimuli

The experiment was an avatar personification (between-subjects manipulation) × intrinsic involvement (measured variable) mixed design. The key manipulation was the personification of the avatar available to assist with the purchase (no avatar, attractive avatar, or expert avatar). We established the attractiveness and expertise of the avatars in three ways: appearance, credentials, and information (product positioning). We purposely aligned these three factors to create a consistent personification for the avatar.[1]

Attractive avatar condition. We designed attractive avatars to appear younger, thinner, and more athletic (see Figure 3). The avatar introduced itself by saying, "Hello, my name is Kim (Tom). I have recently been employed as a shoe consultant. I have already gained preliminary knowledge about our popular, customizable shoe products. I can certainly help you find an attractive offer."[2] The avatar then began the consultation using a fashion-oriented statement that was consistent with the avatar's personification:

> Did you know a person's outfit is judged, above all, by the shoes that he or she is wearing? And comfortable sports shoes are not seen as very stylish. It is too bad that most fashionable shoes look trendy but do not offer much comfort. Using our company, you can design the look of your shoe and make it exactly as you wish. In addition, we can custom-fit your size and design the shoe's features according to your needs.

The avatar was present on the remaining screens to assist in the purchase. When appropriate, we tailored the avatar's advice to be consistent with its personification (for a sample consultation from the attractive avatar condition, see Figure 3).

Expert avatar condition. We designed expert avatars to appear older and nonathletic, and they wore eyeglasses. The avatar introduced itself by saying, "Hello, my name is Dr. Anne Schneider (Dr. Norbert Oswald). I have been a podiatrist for over 10 years and I would like to competently inform and advise you about our products. Take advantage of my experience and let me show you how our ergonomic shoes can be useful to you." The avatar then began the consultation using a comfort-oriented statement that was consistent with the avatar's personification:

> Every day, your feet are subjected to a great deal of strain, which all too often leads to discomfort. Fashionable shoes, unfortunately, are usually ergonomically unsatisfactory, and comfortable shoes often do not have the right look. Our medically tested shoes, on the other hand, can be specially made according to your needs regarding fit and features. Our shoes offer a healthy form and comfort. You can also choose your desired design.

The avatar was present on the remaining screens, and we tailored its advice to be consistent with its personification.

No-avatar control. The no-avatar control condition did not contain an avatar; thus, there was no introductory salutation. There was an attempt to keep the consultation in the no-avatar control equivalent to the consultation in the avatar conditions. We accomplished this by emphasizing both fashion and comfort at the beginning of the consultation:

> We are a company that offers customized leisure shoes. We will introduce our product idea on the following Web pages. When buying leisure shoes, customers can usually choose between trendy fashion shoes and comfortable sport shoes. Our shoes, however, can be individually made for each customer with regard to fit, features, and design.

The information provided during the consultation was a blend of fashion and comfort information.

Procedure

We conducted the online experiment using a population of German shoppers. We recruited the participants using various online and offline procedures. The final sample size was 400 consumers, 55% of whom were men and 45% of whom were women. The participants were between ages 17 and 74 years, and the median age was 24 years. Sixty-six percent of the participants were college graduates, implying a disproportionately high level of education relative to a representative sample of online consumers. In addition, the participants' experience with the Internet and online shopping was above average compared with published statistics about the average Internet user.

We began by recording demographic information. Next, we asked participants four questions about their use of sales associates when shopping. Then, participants saw an introduction to the experiment:

> On the following pages, a company that sells leisure shoes will introduce its products to you. The unique feature of the shoes is that they can be customized with respect to fit, features, and design. Please be aware that these are not actual Web pages for an online store, but rather pages that have been specifically designed for this experiment. You can use these pages to learn about design options, to have an online consultation, and to calculate a price for your personalized shoes.

After the introduction, participants were greeted by the avatar if they were in an avatar condition, received the introductory statement consistent with their condition, and were taken through the customized shoe-shopping experience. During the customization process, the participant could select different visual and functional features to create a personalized shoe. Participants were asked to select a procedure for fitting, the outside shoe material, the inside shoe material, and the soles. The avatar consultant discussed how to make a specific design decision and provided a recommendation before each decision. After the consultation, a randomly generated price of $92.40, $94.40, or $96.40 was provided. Within an avatar condition, the avatar used the same text and made the same recommendations for each participant. The participants were not required to follow the avatar's recommendations. The participants had the opportunity to recustomize the shoe at the end of the process

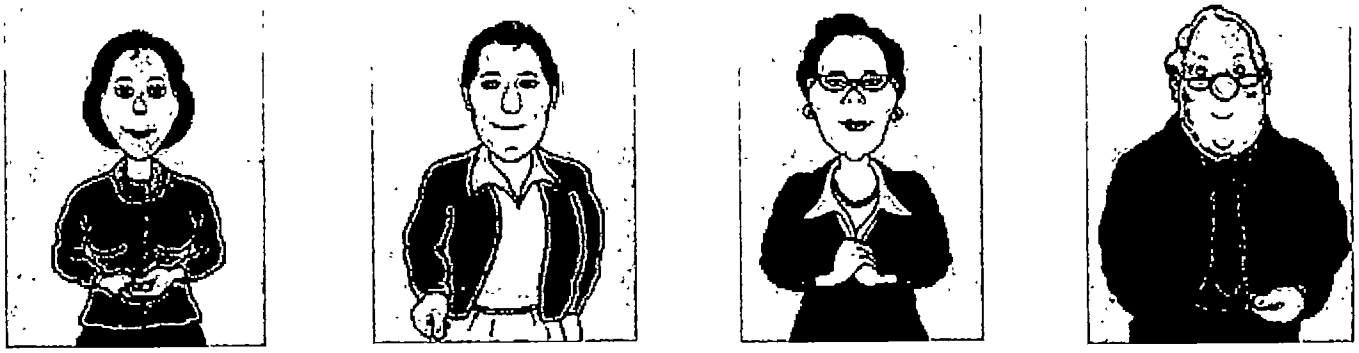
---

[1] In this design, we intentionally confound the avatar's appearance and credentials and the content of the consultation to create an avatar personification factor. Thus, we cannot assess the independent influence of an avatar's presence. In a follow-up study, we control for the content of the consultation and show the independent influence of the avatar.

[2] We included an avatar gender manipulation in the design because industry contacts believed that avatar gender was a relevant factor. Avatar gender did not exhibit a main effect or interact with the manipulated variables for any dependent measure. The stimulus materials were in German. The statements are translations of German text.

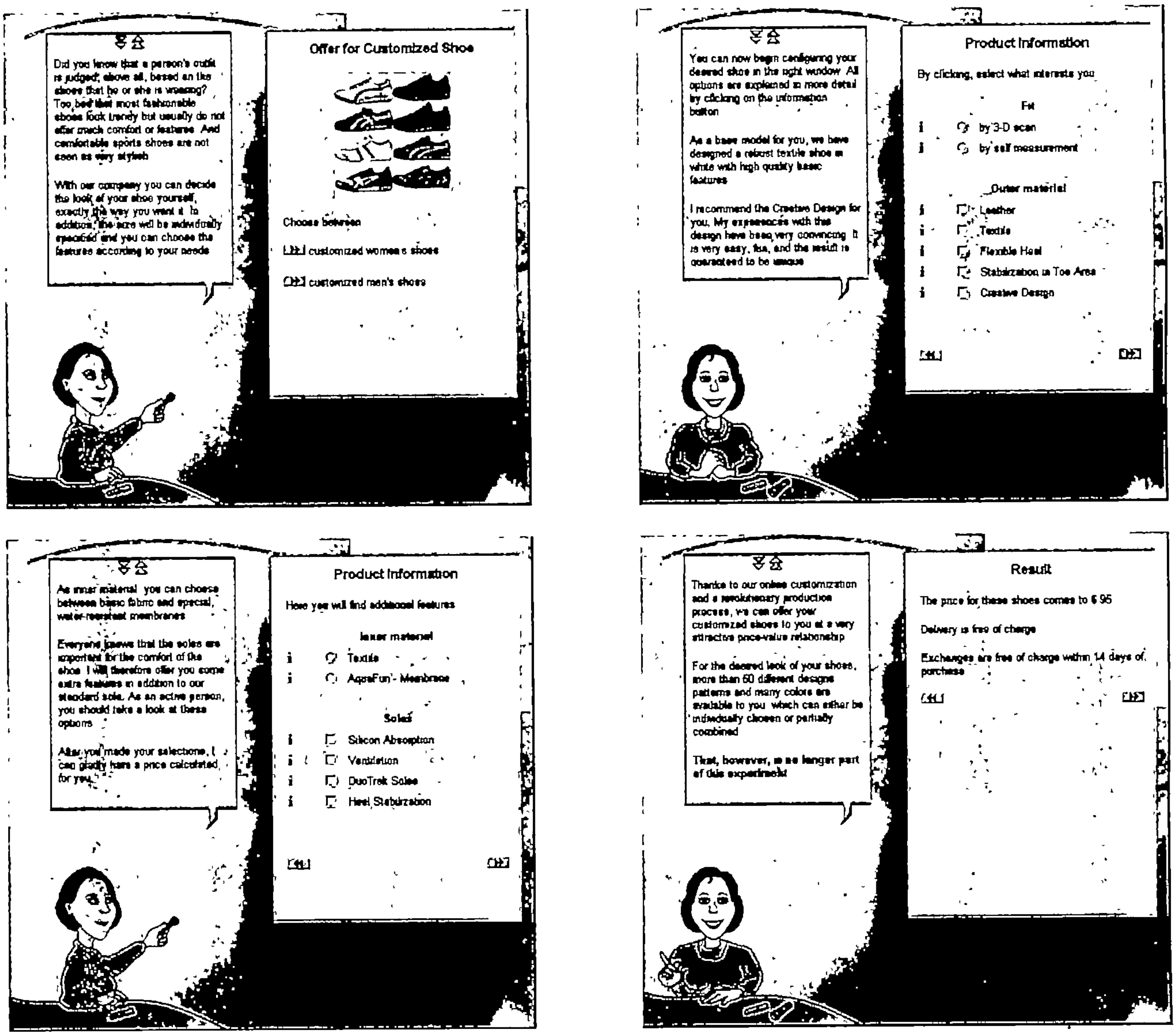
FIGURE 3
Study 1: Avatars and Sample Consultation

A: Avatar Types



Attractive Attractive Expert Expert

B: Consultation Using Appearance-Oriented Product Positioning Text



Did you know that a person's outfit is judged, above all, based on the shoes that he or she is wearing? Too bad that most fashionable shoes look trendy but usually do not offer much comfort or features. And comfortable sports shoes are not seen as very stylish.

With our company you can decide the look of your shoe yourself, exactly the way you want it. In addition, the size will be individually specified and you can choose the features according to your needs.

Offer for Customized Shoe

Choose between

customized women's shoes

customized men's shoes

You can now begin configuring your desired shoe in the right window. All options are explained in more detail by clicking on the information button.

As a base model for you, we have designed a robust textile shoe in white with high quality basic features.

I recommend the Creative Design for you. My experiences with this design have been very convincing. It is very easy, fun, and the result is guaranteed to be unique.

Product Information

By clicking, select what interests you.

Fit
- i ○ by 3-D scan
- i ○ by self measurement

Outer material
- i ☐ Leather
- i ☐ Textile
- i ☐ Flexible Heel
- i ☐ Stabilization in Toe Area
- i ☐ Creative Design

As inner material you can choose between basic fabric and special, water-resistant membranes.

Everyone knows that the soles are important for the comfort of the shoe. I will therefore offer you some extra features in addition to our standard sole. As an active person, you should take a look at these options.

After you made your selections, I can gladly have a price calculated for you.

Product Information

Here you will find additional features

Inner material
- i ○ Textile
- i ○ AquaFun'- Membrane

Soles
- i ☐ Silicon Absorption
- i ☐ Ventilation
- i ☐ DuoTrek Soles
- i ☐ Heel Stabilization

Thanks to our online customization and a revolutionary production process, we can offer your customized shoes to you at a very attractive price-value relationship.

For the desired look of your shoes, more than 60 different designs patterns and many colors are available to you which can either be individually chosen or partially combined.

That, however, is no longer part of this experiment.

Result

The price for these shoes comes to € 95

Delivery is free of charge

Exchanges are free of charge within 14 days of purchase (thus the random assignment of price), though few chose to do so. The participants were not asked to make an actual purchase.

Participants then responded to a series of dependent measures. Participants reported their involvement with the shoe purchase, perceptions of avatar attractiveness and expertise, perceptions of the Web site's entertainment value and information value, perceptions of the Web site's graphic design, perceptions of avatar likeability and credibility, perceptions of retailer credibility, satisfaction with the retailer, attitude toward the product, and purchase intention. We measured all items with seven-point Likert scales, anchored by "completely disagree" and "completely agree." The individual items used to measure each construct and the Cronbach's alpha for each scale appear in Appendix A.

Tests of H_1–H_3

H_1: Avatar effects on satisfaction, attitude, and purchase intention. The means of the satisfaction-with-the-retailer, attitude-toward-the-product, and purchase-intention dependent variables appear in Table 1.[3] As H_1 predicted, participants who saw an avatar were more satisfied with the retailer (M = 3.91) than those who did not see an avatar (M = 3.41; F(1, 397) = 9.19, $p < .01$), had a more favorable attitude toward the product (M = 4.36) than those who did not see an avatar (M = 3.86; F(1, 397) = 7.31, $p < .01$), and had a higher purchase intention (M = 3.74) than those who did not see an avatar (M = 3.00; F(1, 397) = 9.16, $p < .01$).

H_2: Avatar effects on perception of the Web site. H_2 predicted that avatars would facilitate Web-based purchasing because they make the Web site seem more entertaining and informative. Participants who saw an avatar rated the Web site as more entertaining ($M_{\text{avatar}} = 4.22$, $M_{\text{no avatar}} = 3.55$; F(1, 397) = 13.45, $p < .01$) and informative ($M_{\text{avatar}} = 4.48$, $M_{\text{no avatar}} = 3.99$; F(1, 397) = 9.12, $p < .01$) than participants who did not see an avatar.

H_3: The mediating effects of Web site perception. H_3 predicted that the increase in the perceived entertainment value and information value of the Web site would mediate the influence of the avatar's presence on satisfaction with the retailer, attitude toward the product, and purchase intention. Evidence for complete mediation requires (1) an effect of the independent variable (i.e., avatar presence) on the dependent variable (i.e., satisfaction with the retailer, attitude toward the product, and purchase intention), (2) effects of the independent variable (i.e., avatar presence) on the mediators (i.e., entertainment, informativeness), and (3) a nonsignificant effect of the independent variable when the mediators and independent variable are regressed on the dependent variable (Baron and Kenny 1986). As we summarize in Table 2, the presence of an avatar significantly influenced satisfaction with the retailer (t(398) = 3.03, $p < .05$), attitude toward the product (t(398) = 2.71, $p < .05$), and purchase intention (t(398) = 3.03, $p < .05$). The presence of the avatar also influenced the entertainment value (t(398) = 3.67, $p < .05$) and the informativeness (t(398) = 3.02, $p < .05$) of the Web site. Finally, the inclusion of entertainment value and avatar presence in a regression made the avatar's presence a nonsignificant predictor of satisfaction with the retailer (t(397) = 1.16, $p > .05$), attitude toward the product (t(397) = 1.14, $p > .05$), and purchase intention (t(397) = 1.52, $p > .05$). Likewise, the inclusion of informativeness and avatar presence in a regression made the avatar's presence a nonsignificant predictor of satisfaction with the retailer (t(397) = 1.39, $p > .05$), attitude toward the product (t(397) = 1.38, $p > .05$), and purchase intention (t(397) = 1.80, $p > .05$). These tests provide evidence for mediation (Baron and Kenny 1986).

Tests of H_4 and H_5

We used the 322 participants in the attractive and expert avatar conditions to test H_4 and H_5. The hypotheses address whether participants with different levels of involvement with the product being purchased were more responsive to an attractive or expert avatar. Given the continuous nature of the intrinsic involvement variable, we used ordinary least squares regression to test the hypotheses. We conducted follow-up tests of avatar effectiveness at different levels of involvement using an analysis of variance (ANOVA). We

TABLE 1
Study 1: Means

	Avatar Condition		
Dependent Measure	**None (n = 78)**	**Attractive (n = 159)**	**Expert (n = 163)**
Satisfaction with retailer	3.41	3.93	3.89
Attitude toward product	3.86	4.38	4.34
Purchase intention	3.00	3.86	3.60
Web site entertainment	3.55	4.28	4.16
Web site informativeness	3.99	4.44	4.51
Avatar attractiveness		2.91	2.29
Avatar expertise		4.15	4.72
Avatar likeability		4.24	4.09
Avatar credibility		3.78	4.34

---

[3] The smaller cell sizes for the control group are a consequence of a communication problem with the Web site programmer. The programmer considered gender an experimental rather than a counterbalance factor; thus, the male avatar, female avatar, and control conditions have similar cell sizes.

TABLE 2
Study 1: Mediation Analysis for Effects of Avatar Presence

Test	Variables	t	p
A → C	Avatar → satisfaction with retailer	3.03	<.01
	Avatar → attitude toward product	2.71	<.01
	Avatar → purchase intention	3.03	<.01
A → B	Avatar → entertainment	3.67	<.01
	Avatar → informativeness	3.02	<.01
Mediation Tests for Entertainment			
A, B → C	Entertainment,	13.53	<.01
	Avatar → satisfaction with retailer	1.16	.25
	Entertainment,	10.25	<.01
	Avatar → attitude toward product	1.14	.26
	Entertainment,	10.01	<.01
	Avatar → purchase intention	1.52	.13
Mediation Tests for Informativeness			
A, B → C	Informativeness,	17.05	<.01
	Avatar → satisfaction with retailer	1.39	.17
	Informativeness,	11.32	<.01
	Avatar → attitude toward product	1.38	.17
	Informativeness,	10.48	<.01
	Avatar → purchase intention	1.80	.07

created the levels of involvement using a tertiary split of the self-reported involvement scores.

Manipulation check. We performed two types of manipulation checks. First, we confirmed that the avatar manipulations were effective. Participants perceived the attractive avatar (M = 2.91) as more attractive than the expert avatar (M = 2.29; F(1, 319) = 17.55, $p < .05$), and they perceived the expert avatar (M = 4.72) as more expert than the attractive avatar (M = 4.15; F(1, 319) = 15.16, $p < .05$). Second, we confirmed that involvement and type of avatar were orthogonal factors and thus could be used in the same regression analysis. The avatar manipulation did not influence the degree of involvement with the product ($M_{\text{no avatar}} = 3.82$, $M_{\text{attractive}} = 3.65$, $M_{\text{expert}} = 3.60$; F(2, 397) = .73, $p > .05$). In addition, the degree of involvement with the product did not interact with perceptions of the attractiveness of the attractive avatar (F(1, 318) = .28, $p > .05$) or perceptions of the expertise of the expert avatar (F(1, 318) = .15, $p > .05$).[4]

Review of hypotheses. We offered two general hypotheses about the effectiveness of different types of avatars. First, we predicted that the more attractive avatar would appeal to participants who were moderately interested in custom-fit shoes (H_{4a}). The reason for this preference is that participants would perceive the avatar as more likeable (H_{5a}). Second, we predicted that the more expert avatar would appeal to participants who were highly interested in custom-fit shoes (H_{4b}). The reason for this preference is that participants would perceive the avatar as more credible (H_{5b}). This pattern of results required participants to be more sensitive to avatar attractiveness as involvement increased to a moderate level and more sensitive to avatar expertise as involvement increased to a high level. As we stated in H_4, the response to the attractive avatar needed to show a quadratic form with increasing involvement, whereas the response to the expert avatar needed to show a linear form with increasing involvement (see Figure 2).

H_4: Satisfaction with the retailer. The first analysis investigated the type of avatar × involvement interaction for satisfaction with the retailer. Consistent with H_4, there was a type of avatar × involvement interaction with a quadratic influence of involvement across the attractive and expert conditions (F(1, 316) = 10.16, $p < .05$).[5] A review of the

---

[4] The test of the type of avatar × involvement interaction involved a comparison of a constrained model ($\hat{Y} = \alpha + \beta \times \text{TYPE} + \beta \times \text{INVOLVE}$) with an unconstrained model ($\hat{Y} = \alpha + \beta \times \text{TYPE} + \beta \times \text{INVOLVE1} + \beta \times \text{INVOLVE2}$), where TYPE was dummy-coded 0/1, INVOLVE was the involvement score, INVOLVE1 was the involvement scores for the attractive treatment group (zero otherwise), and INVOLVE2 was the involvement scores for the expert treatment group (zero otherwise).

[5] The test of the type of avatar × involvement interaction involved a comparison of a constrained model that required a natural log for involvement ($\hat{Y} = \alpha + \beta \times \text{TYPE} + \beta \times \text{INVOLVE1} + \beta \times \text{INVOLVE2} + \beta \times \ln[\text{INVOLVE}]$) with a less constrained model ($\hat{Y} = \alpha + \beta \times \text{TYPE} + \beta \times \text{INVOLVE1} + \beta \times \text{INVOLVE2} + \beta \times \ln(\text{INVOLVE1}) + \beta \times \ln[\text{INVOLVE2}]$), where TYPE was dummy-coded 0/1, INVOLVE was the involvement score, INVOLVE1 was the involvement scores for the attractive group (zero otherwise), and INVOLVE2 was the involvement scores for the expert group (zero otherwise). Note that there was no type of avatar × involvement interaction with a linear influence of involvement across the attractive and expert conditions (F(1, 318) = .32, $p > .05$).

beta coefficients showed a nonsignificant influence of the linear involvement term (t(316) = –1.13, $p > .05$) and a significant influence of the quadratic involvement term (t(316) = –2.85, $p < .05$) in the attractive avatar condition. There was a significant influence of the linear involvement term (t(316) = 3.62, $p < .05$) and a nonsignificant influence of the quadratic involvement term (t(316) = –1.58, $p > .05$) in the expert avatar condition.

To help illustrate this interaction, the data are shown using a tertiary split on the involvement independent variable (see Table 3). The data show that moderately involved participants were more satisfied with the retailer when they shopped using the attractive avatar, but highly involved participants were more satisfied when they shopped using the expert avatar. Although the regression analysis could not test these predictions at various levels of involvement, the tertiary grouping of participants by involvement level allowed for an ANOVA of simple effect tests, albeit with less power to detect a difference between the avatar conditions (MacCallum et al. 2002). The attractive avatar (M = 3.10) and expert avatar (M = 3.11) resulted in equivalent levels of satisfaction at low levels of involvement (F(1, 316) = .01, $p > .05$). The attractive avatar (M = 4.35) led to more satisfaction than the expert avatar (M = 3.92) for moderately involved participants (F(1, 316) = 3.45, $p < .05$).[6] The expert avatar (M = 4.85) did not lead to more satisfaction than the attractive avatar (M = 4.55) for highly involved participants (F(1, 316) = 1.86, $p > .05$), though the means were consistent with predictions. In summary, the regression analysis provided support for H_4, and the ANOVA provided support for H_{4a} but only directional support for H_{4b}.

H_4: Attitude toward the product. Consistent with H_4, there was a type of avatar × involvement interaction for attitude toward the product with a quadratic influence of involvement across the attractive and expert conditions (F(1, 316) = 5.35, $p < .05$).[7] The beta coefficients showed a nonsignificant influence of the linear involvement term (t(316) = –.58, $p > .05$) and a significant influence of the quadratic involvement term (t(316) = 2.83, $p < .05$) in the attractive avatar condition. There was a significant influence of the linear involvement term (t(316) = 3.62, $p < .05$) and a nonsignificant influence of the quadratic involvement term (t(316) = –.50, $p > .05$) in the expert avatar condition.

Using the tertiary split on the involvement independent variable (see Table 3), the attractive avatar (M = 3.37) and the expert avatar (M = 3.43) resulted in an equivalent attitude toward the product at low levels of involvement (F(1, 316) = .10, $p > .05$). The attractive avatar (M = 4.68) led to a more positive attitude toward the product than the expert avatar (M = 4.22) for moderately involved participants (F(1, 316) = 3.78, $p < .05$). The expert avatar (M = 5.64) led to a marginally more positive attitude than the attractive avatar (M = 5.28) for highly involved participants (F(1, 316) = 2.41, $p = .06$). In summary, the regression analysis provided support for H_4, and the ANOVA provided support for H_{4a} and marginal support for H_{4b}.

H_4: Purchase intention. Consistent with H_4, there was a type of avatar × involvement interaction for purchase intention with a unique quadratic influence of involvement in the attractive and expert conditions (F(1, 316) = 5.94, $p < .05$).[8] The beta coefficients showed a nonsignificant influence of

TABLE 3
Study 1: Means by Level of Involvement

Dependent Variable and Condition	n	Avatar None	 Attractive	 Expert
Satisfaction with Retailer				
Low involvement	143	2.37	3.10	3.11
Moderate involvement	123	3.31	4.35	3.92
High involvement	134	4.16	4.55	4.85
All	400	3.41	3.93	3.89
Attitude Toward Product				
Low involvement	143	2.51	3.37	3.43
Moderate involvement	123	3.95	4.68	4.22
High involvement	134	4.68	5.28	5.64
All	400	3.86	4.38	4.34
Purchase Intention				
Low involvement	143	1.44	2.52	2.59
Moderate involvement	123	2.91	4.15	3.48
High involvement	134	4.07	4.61	5.02
All	400	3.00	3.86	3.60

---

[6] Given the directional hypotheses in the moderate- and high-involvement conditions, we used one-tailed tests.

[7] Note that there was no type of avatar × involvement interaction for attitude toward the product with a linear influence of involvement across the attractive and expert conditions (F(1, 318) = .19, $p > .05$).

[8] Note that there was no type of avatar × involvement interaction for purchase intention with a linear influence of involvement across the attractive and expert conditions (F(1, 318) = .95, $p > .05$).

the linear involvement term ($t(316) = -.09$, $p > .05$) and a significant influence of the quadratic involvement term ($t(316) = 2.01$, $p < .05$) in the attractive avatar condition. There was a significant influence of the linear involvement term ($t(316) = 3.76$, $p < .05$) and a nonsignificant influence of the quadratic involvement term ($t(316) = -1.41$, $p > .05$) in the expert avatar condition.

Using the tertiary split on the involvement independent variable (see Table 3), the attractive avatar ($M = 2.51$) and the expert avatar ($M = 2.59$) resulted in an equivalent purchase intention at low levels of involvement ($F(1, 316) = .07$, $p > .05$). The attractive avatar ($M = 5.15$) led to more positive purchase intentions than the expert avatar ($M = 3.41$) for moderately involved participants ($F(1, 316) = 5.46$, $p < .05$). The expert avatar ($M = 5.02$) led to a marginally more positive purchase intention than the attractive avatar ($M = 4.61$) for highly involved participants ($F(1, 316) = 2.21$, $p = .07$). Again, the regression analysis provided support for H_4, and the ANOVA provided support for H_{4a} and marginal support for H_{4b}.

H_{5a}. H_{5a} predicted that avatar likeability would mediate the persuasiveness of the attractive avatar relative to the expert avatar. As the test of H_{4a} shows, the attractive avatar was more persuasive than the expert avatar at moderate levels of involvement. Thus, we used the 97 participants who viewed an avatar and were moderately involved in the product category in the mediation analysis. As we summarize in Table 4, the presence of an attractive, as opposed to an expert, avatar significantly influenced satisfaction with the retailer ($t(96) = 1.85$, $p = .07$), attitude toward the product ($t(96) = 2.12$, $p < .05$), and purchase intention ($t(96) = 2.16$, $p < .05$). The attractive avatar was more liked ($t(96) = 1.83$, $p = .07$) but was not more credible ($t(96) = 1.34$, $p > .05$) than the expert avatar. Finally, the inclusion of likeability and type of avatar in a regression made the type of avatar a nonsignificant predictor of satisfaction with the retailer ($t(95) = 1.11$, $p > .05$), attitude toward the product ($t(95) = 1.54$, $p > .05$), and purchase intention ($t(96) = 1.62$, $p > .05$). Sobel tests showed that avatar likeability was only a partial mediator of the type of avatar effect. Table 3 also reports nonsignificant tests of avatar credibility as a mediator, as we expected.

H_{5b}. H_{5b} predicted that avatar credibility would mediate the persuasiveness of the expert avatar relative to the attractive avatar. As the test of H_{4b} shows, the expert avatar was marginally more persuasive than the attractive avatar at high levels of involvement. Thus, we used the 101 participants who viewed an avatar and were highly involved in the product category in the mediation analysis. As we summarize in Table 3, the presence of an expert, as opposed to an attractive, avatar did not influence satisfaction with the retailer ($t(100) = 1.56$, $p = .12$), influenced attitude toward the product ($t(100) = 1.96$, $p < .05$), and marginally influenced purchase intention ($t(100) = 1.69$, $p < .05$). The expert avatar was more credible ($t(100) = 3.96$, $p < .05$) but was not more liked ($t(100) = .79$, $p > .05$) than the attractive avatar. Finally, the inclusion of credibility and type of avatar in a regression made the type of avatar a nonsignificant predictor of satisfaction with the retailer ($t(99) = .30$, $p > .05$), attitude toward the product ($t(99) = .67$, $p > .05$), and purchase intention ($t(99) = .97$, $p > .05$). Sobel tests showed that avatar credibility was only a partial mediator of the type of avatar effect. Table 3 also reports nonsignificant tests of avatar likeability as a mediator, as we expected.

TABLE 4
Study 1: Mediation Analysis for Effects of Avatar Type

		Moderate Involvement			High Involvement		
Test	**Variables**	**Predicted**	**t**	***p***	**Predicted**	**t**	***p***
A → C	Type of avatar → satisfaction with retailer	Significant	1.85	.07	Significant	1.56	.12
	Type of avatar → attitude toward product	Significant	2.12	.04	Significant	1.96	.05
	Type of avatar → purchase intention	Significant	2.16	.03	Significant	1.69	.09
A → B	Type of avatar → likeability	Significant	1.83	.07	n.s.	.79	.43
	Type of avatar → credibility	n.s.	1.34	.17	Significant	3.96	.00
Mediation Tests for Likeability							
A, B → C	Likeability,	Significant	5.09	.00	None	4.39	.00
	Type of avatar → satisfaction with retailer	n.s.	1.11	.27	Significant	1.35	.18
	Likeability,	Significant	3.60	.01	None	2.90	.01
	Type of avatar → attitude toward product	n.s.	1.54	.13	Significant	1.79	.08
	Likeability,	Significant	3.27	.00	None	1.72	.09
	Type of avatar → purchase intention	n.s.	1.62	.11	Significant	1.59	17
Mediation Tests for Credibility							
A, B → C	Credibility,	None	5.33	.00	Significant	3.34	.00
	Type of avatar → satisfaction with retailer	Significant	2.83	.01	n.s.	.30	.76
	Credibility,	None	3.02	.00	Significant	3.37	.00
	Type of avatar → attitude toward product	Significant	2.61	.01	n.s.	.67	.54
	Credibility,	None	2.97	.00	Significant	1 69	.10
	Type of avatar → purchase intention	Significant	2.64	.01	n.s.	.97	.33

Notes: n.s. = not significant.

Discussion

Study 1 provides three compelling results. First, the simple inclusion of an avatar on the screens of a Web-based shopping site increased consumer satisfaction with the retailer, made the customer's attitude toward the product more favorable, and increased the customer's purchase intention (H_1). Second, the perceived entertainment value and informativeness of the Web site mediated the influence of the avatars on the key dependent measures (H_2 and H_3). Third, the responsiveness to a particular type of avatar depended on the type of avatar and the consumer's involvement with the product category (H_4). Attractive avatars were more persuasive when consumers were moderately involved in the purchase (H_{4a}), whereas expert avatars were marginally more persuasive when consumers were highly involved in the purchase (H_{4b}). These results are consistent with many empirical results showing that differential characteristics of persuasion agents are effective at different levels of customer involvement with the purchase.

There are two unresolved issues in the results of Study 1. First, our effort to personify the avatars using an externally valid procedure allows for an alternative explanation of the avatar effects (e.g., H_1–H_3) observed in Study 1. It may be that the consultation in the no-avatar condition was of poorer quality than the consultations in the avatar conditions. Recall that the consultation in the no-avatar condition was a combination of fashion and comfort statements, whereas the consultation in the attractive (expert) avatar condition was fashion oriented (comfort oriented). To address this problem, we replicated Study 1 using an avatar type (no avatar, attractive avatar, or expert avatar) × consultation type (fashion, comfort, or fashion and comfort) design. We randomly assigned 659 participants to one of the nine conditions. The results of this follow-up study replicated all the avatar-presence results (i.e., H_1 and H_2) and the mediation analysis results (H_3) of Study 1.[9] More important, there was no interaction of the consultation type with the avatar type on any of the dependent measures. Thus, it appears that the avatar and its credentials, not the content of the consultation, are responsible for the effects we observed in Study 1.

The second issue pertains to the source of the involvement effects (H_4 and H_5) observed in Study 1. The data show that attractive avatars are more persuasive than expert avatars at moderate levels of involvement and that expert avatars are more persuasive than attractive avatars at high levels of involvement (H_{4a} and H_{4b}). Yet likeability only partially mediates the type of avatar effect at moderate levels of involvement, and credibility only partially mediates the type of avatar effect at high levels of involvement (H_5). We expect that there are two reasons for partial mediation. First, the effects of attractiveness and expertise at their respective levels of involvement are modest (i.e., *p* values range from .03 to .12). If a relationship between an independent and a dependent variable is modest, the ability of a mediator to explain this relationship is limited. A solution to this problem may be to increase the statistical power of the test. Power ranged from .32 to .57 in the tests of H_{4a} and H_{4b}, well below the recommended levels of .80. Second, the manipulations of attractiveness and expertise were not independent in Study 1. Although this did not create a problem for testing H_1–H_3, in which we treated avatars as replicates, it may have caused a problem for testing H_5. For example, at moderate levels of involvement, participants perceived an attractive avatar as more attractive than an expert avatar but also as less credible than an expert avatar. This means that the likeability score for the attractive avatar could be influenced in opposite directions by attractiveness and expertise (i.e., the likeability score could be biased downward). Consequently, the mediation tests involving likeability could have been compromised. A similar phenomenon could have biased the credibility ratings of the expert avatar, compromising the mediation tests involving credibility. Study 2 investigates H_4 and H_5 using an experimental design that controls for these potential biases.

Study 2

Study 2 investigated the influence of avatar attractiveness and avatar expertise on persuasion at different levels of intrinsic involvement with a product. The experiment used a level-of-attractiveness (low and high) × level-of-expertise (low and high) factorial design along with a measure of the consumer's intrinsic involvement with the product. We expected that increases in avatar attractiveness would be most effective for moderately involved consumers and that the degree of persuasion would be mediated by the likeability of the avatar. Similarly, we expected that increases in avatar expertise would be most effective for highly involved consumers and that the degree of persuasion would be mediated by the credibility of the avatar.

Stimuli and Procedure

The stimuli and procedure were similar to Study 1. The primary changes involved orthogonal manipulations of avatar attractiveness and avatar expertise. For example, one possibility was to attempt to draw avatars that would be perceived as low on both factors, high on both factors, or a mix of levels across the factors. Pretests revealed that it was difficult to draw avatars that unambiguously fit these four profiles. Thus, we decided to manipulate avatar attractiveness using the avatar's visual features and avatar expertise using the avatar's credentials and style of consultation. First, we redesigned the female avatars to create a stronger manipulation of attractiveness.[10] Second, we rewrote the consultation to be consistent with a low-expertise or high-expertise consultant (see Appendix B). In summary, the experiment was an attractiveness (low and high) × expertise (low and high) between-subjects design. In all other respects, the procedure was identical to Study 1.

---

[9] The results are available from the second author.

[10] We used only female avatars in this study because of (1) the lack of a gender effect in Study 1 and (2) the expense associated with redesigning the avatars.

University Library Calcutta

Results

We recruited participants from a population of German online shoppers. The final sample size was 596 consumers, 48.2% of whom were men and 51.8% of whom were women. The participants were between 14 and 80 years of age, and the median age was 28 years.[11] Seventy-four percent of the participants were college graduates, and the participants' experience with online shopping was above average compared with the average Internet user.

Manipulation check. We performed two types of manipulation checks. First, we confirmed that the manipulations were effective. Participants perceived the attractive avatar (M = 4.07) as more attractive than the unattractive avatar (M = 2.06; F(1, 592) = 465.05, $p < .01$), and they perceived the expert avatar (M = 4.65) as more expert than the nonexpert avatar (M = 3.43; F(1, 592) = 118.23, $p < .01$). Second, we confirmed that involvement was orthogonal to the experimental manipulations; thus, both factors could be used in the same regression analysis. The avatar attractiveness manipulation influenced the degree of involvement with the product ($M_{attractive}$ = 3.54, $M_{unattractive}$ = 3.29; F(1, 592) = 4.57, $p < .05$), but this effect was small compared with the variability in involvement (i.e., less than 2%). The avatar expertise manipulation did not influence the degree of involvement with the product (M_{expert} = 3.41, $M_{nonexpert}$ = 3.42; F(1, 592) = .04, $p > .05$).

Review of hypotheses. We offered two general hypotheses about the effectiveness of different types of avatars. First, we expected that the more attractive avatar would appeal to participants who were moderately interested in custom-fit shoes (H_{4a}). The reason for this preference is that participants would perceive the avatar as more likeable (H_{5a}). Second, we expected that the more expert avatar would appeal to participants who were highly interested in custom-fit shoes (H_{4b}). The reason for this preference is that participants would perceive the avatar as more credible (H_{5b}).

Initial analyses. Similar to Study 1, the initial analysis used regressions to test H_4. We anticipated that the impact of the attractiveness manipulation would be greater for moderately involved consumers than for mildly or strongly involved consumers. Although the means were consistent with this prediction (see Table 5), tests for a quadratic influence of involvement across the two attractiveness conditions were not significant for satisfaction with the retailer (t(591) = .76, $p > .05$), attitude toward the product (t(591) = 1.02, $p > .05$), or purchase intention (t(591) = .56, $p > .05$).[12] In general, the attractiveness manipulation was so strong that it had an influence at all levels of involvement. We anticipated that the impact of the expertise manipulation would become greater as involvement increased. Tests for an expertise × involvement interaction were significant for satisfaction with the retailer (t(591) = 5.29, $p < .01$), attitude toward the product (t(591) = 3.70, $p < .01$), and purchase intention (t(591) = 3.87, $p < .01$).

H_{4a}: Moderate involvement. H_{4a} predicted that at moderate levels of involvement, there would be a main effect for the attractiveness manipulation but not for the expertise manipulation. The data are shown using a tertiary split on the involvement independent variable (see Table 5). People who viewed an attractive avatar were more satisfied with the retailer ($M_{attractive}$ = 3.75 [e.g., Ms = 3.54 and 3.96], $M_{unattractive}$ = 3.13; F(1, 590) = 18.58, $p < .01$), had a more positive attitude toward the product ($M_{attractive}$ = 4.53, $M_{unattractive}$ = 3.90; F(1, 590) = 18.57, $p < .01$), and had a more positive purchase intention ($M_{attractive}$ = 3.58,

---

[11] The ten shoppers, ages 14–16, had Internet shopping experience, so we retained them in the sample.

[12] Note that the attractiveness manipulation shows a log relationship to involvement in the attractive–nonexpert condition. This condition is the conceptual equivalent of the attractive avatar condition in Study 1.

TABLE 5
Study 2: Means by Level of Involvement

Dependent Variable and Condition	n	Unattractive Nonexpert	Unattractive Expert	Attractive Nonexpert	Attractive Expert
Satisfaction with Retailer					
Low involvement	201	2.60	2.60	3.03	2.94
Moderate involvement	201	3.21	3.03	3.54	3.96
High involvement	194	3.37	4.23	3.82	4.82
All	596	3.03	3.23	3.51	3.93
Attitude Toward Product					
Low involvement	201	2.76	2.87	3.01	3.41
Moderate involvement	201	3.75	4.08	4.55	4.51
High involvement	194	4.72	5.16	4.86	5.42
All	596	3.67	3.95	4.27	4.47
Purchase Intention					
Low involvement	201	1.78	1.89	1.67	2.12
Moderate involvement	201	3.07	2.82	3.79	3.37
High involvement	194	4.00	4.22	4.09	4.95
All	596	2.87	2.89	3.36	3.53

$M_{unattractive} = 2.95$; $F(1, 590) = 12.74$, $p < .01$). People who viewed an expert avatar were not more satisfied with the retailer ($M_{expert} = 3.52$, $M_{nonexpert} = 3.36$; $F(1, 590) = 1.07$, $p > .05$), did not have a more positive attitude toward the product ($M_{expert} = 4.30$, $M_{nonexpert} = 4.13$; $F(1, 590) = 1.35$, $p > .05$), and did not have a more positive purchase intention ($M_{expert} = 3.11$, $M_{nonexpert} = 3.41$; $F(1, 590) = 2.79$, $p > .05$). The interaction between avatar attractiveness and expertise was not significant for retailer satisfaction ($F(1, 588) = 3.74$, $p > .05$), attitude toward the product ($F(1, 588) = 1.62$, $p > .05$), and purchase intention ($F(1, 588) = .27$, $p > .05$).

H_{5a}: Moderate involvement. H_{5a} predicted that avatar likeability would mediate the persuasiveness of the attractive avatar relative to the unattractive avatar. We used the 201 participants who were moderately involved in the product category in the mediation analysis. As we summarize in Table 6, the presence of an attractive, as opposed to an unattractive, avatar significantly influenced satisfaction with the retailer ($t(200) = 4.31$, $p < .01$), attitude toward the product ($t(200) = 4.31$, $p < .01$), and purchase intention ($t(200) = 3.57$, $p < .01$). The second test showed that the attractive avatar was more liked ($M = 4.73$) than the unattractive avatar ($M = 3.59$; $t(200) = 6.46$, $p < .01$). Finally, the inclusion of likeability and avatar attractiveness in a regression equation made the type of avatar a nonsignificant predictor of satisfaction with the retailer ($t(199) = 1.76$, $p > .05$) and attitude toward the product ($t(199) = 1.82$, $p > .05$), whereas purchase intention remained significant ($t(199) = 2.38$, $p < .05$). Sobel tests showed that avatar likeability was a significant mediator of avatar attractiveness for satisfaction with the retailer ($t(199) = 4.32$, $p < .01$), attitude toward the product ($t(199) = 4.32$, $p < .01$), and purchase intention ($t(199) = 2.05$, $p < .05$).

H_{4b}: High involvement. H_{4b} predicted that at high levels of involvement, there would be a main effect for the expertise manipulation but not for the attractiveness manipulation. The data are shown using a tertiary split on the involvement independent variable (see Table 5). People who viewed an expert avatar were more satisfied with the retailer ($M_{expert} = 4.56$, $M_{nonexpert} = 3.60$; $F(1, 590) = 18.69$, $p < .05$), had a more positive attitude toward the product ($M_{expert} = 5.30$, $M_{nonexpert} = 4.79$; $F(1, 590) = 11.76$, $p < .05$), and had a more positive purchase intention ($M_{expert} = 4.63$, $M_{nonexpert} = 4.04$; $F(1, 590) = 7.08$, $p < .05$). People who viewed an attractive avatar were more satisfied with the retailer ($M_{attractive} = 4.37$, $M_{unattractive} = 3.80$; $F(1, 590) = 12.04$, $p < .01$), did not have a more positive attitude toward the product ($M_{attractive} = 5.17$, $M_{unattractive} = 4.94$; $F(1, 590) = 2.22$, $p > .05$), and had a more positive purchase intention ($M_{attractive} = 4.57$, $M_{unattractive} = 411$; $F(1, 590) = 4.29$, $p < .01$). The interaction between avatar attractiveness and expertise was not significant for retailer satisfaction ($F(1, 588) = .50$, $p > .05$), attitude toward the product ($F(1, 588) = .18$, $p > .05$), and purchase intention ($F(1, 588) = 3.65$, $p > .05$).

H_{5b}: High involvement. H_{5b} predicted that avatar credibility would mediate the persuasiveness of the expert avatar relative to the nonexpert avatar. We used the 194 participants who were highly involved in the product category in the mediation analysis. As we summarize in Table 6, the presence of an expert, as opposed to a nonexpert, avatar sig-

TABLE 6
Study 2: Mediation Analyses

		Moderate Involvement			High Involvement		
Test	**Variables**	**Predicted**	**t**	***p***	**Predicted**	**t**	***p***
Attractiveness Analysis							
A → C	Avatar attractiveness → satisfaction with retailer	Significant	4.31	.00	n.s.	3.47	.00
	Avatar attractiveness → attitude toward product	Significant	4.31	.00	n.s.	1.49	.14
	Avatar attractiveness → purchase intention	Significant	3.57	.01	n.s.	2.06	.04
A → B	Avatar attractiveness → likeability	Significant	6.46	.00	None	5.32	.00
A, B → C	Likeability,	Significant	5.99	.00	None	9.01	.53
	Avatar attractiveness → satisfaction with retailer	n.s.	1.76	.08	None	.63	.00
	Likeability,	Significant	5.81	.00	None	5.21	.00
	Avatar attractiveness → attitude toward product	n.s.	1.82	.07	None	.39	.70
	Likeability,	Significant	2.16	.03	None	4.66	.00
	Avatar attractiveness → purchase intention	n.s.	2.38	.02	None	.35	.74
Expertise Analysis							
A → C	Avatar expertise → satisfaction with retailer	n.s.	1.03	.31	Significant	6.22	.00
	Avatar expertise → attitude toward product	n.s.	1.16	.27	Significant	3.43	.00
	Avatar expertise → purchase intention	n.s.	1.67	.10	Significant	2.66	.00
A → B	Avatar expertise → credibility	None	3.33	.00	Significant	6.89	.00
A, B → C	Credibility,	None	6.81	.00	Significant	7.27	.00
	Avatar expertise → satisfaction with retailer	None	.45	.65	n.s.	3.04	.00
	Credibility,	None	6.06	.00	Significant	5.58	.00
	Avatar expertise → attitude toward product	None	.17	.87	n.s.	.83	.41
	Credibility,	None	2.72	.01	Significant	4.06	.00
	Avatar expertise → purchase intention	None	2.27	.02	n.s.	.68	.50

Notes: n.s. = not significant.

nificantly influenced satisfaction with the retailer (t(193) = 6.22, $p < .01$), attitude toward the product (t(193) = 3.43, $p < .01$), and purchase intention (t(193) = 2.66, $p < .01$). The second test showed that the expert avatar was more credible (M = 4.65) than the nonexpert avatar (M = 3.37; t(193) = 6.89, $p < .01$). Finally, the inclusion of expertise and avatar credibility in a regression equation made the type of avatar a nonsignificant predictor of attitude toward the product (t(192) = .83, $p > .05$) and purchase intention (t(192) = .68, $p < .05$), whereas satisfaction with the retailer remained significant (t(193) = 3.04, $p > .05$). Sobel tests showed that avatar credibility was a significant mediator of avatar expertise for satisfaction with the retailer (t(193) = 5.01, $p < .01$), attitude toward the product (t(193) = 4.34, $p < .01$), and purchase intention (t(193) = 3.51, $p < .01$).

Discussion

In general, the results of Study 2 are consistent with H_4 and H_5. At moderate levels of involvement with the product purchase, participants were sensitive to the attractiveness of the avatar but not to the expertise of the avatar. The likeability of the avatar mediated the influence of avatar attractiveness on satisfaction with the retailer, attitude toward the product, and purchase intention. At high levels of involvement with the product purchase, participants were sensitive to the expertise of the avatar. The perceived credibility of the avatar mediated the influence of avatar expertise on satisfaction with the retailer, attitude toward the product, and purchase intention. Unexpectedly, there was a consequence of avatar attractiveness at high levels of involvement. This effect may be an influence of the strong manipulation of avatar attractiveness.

General Discussion

The two studies suggest that avatars can enhance the effectiveness of a Web-based sales channel. Study 1 showed that adding an avatar to Web-based information increased the customer's satisfaction with the retailer, attitude toward the product, and purchase intention. A follow-up study showed that the response to the avatar was robust across different types of information content. The avatar itself, not the information provided by the avatar, was the key to persuasion. Study 2 showed that active attempts to manipulate the attractiveness and expertise of the avatar made the avatar more persuasive for certain segments of shoppers. In general, making the avatar more attractive was effective across all levels of involvement. Making the avatar more expert was effective only at high levels of involvement. Study 2 also provided evidence of mediation. The attractiveness of the avatar influenced perceptions of likeability, and likeability mediated the influence of the avatar's attractiveness on persuasion. Similarly, the expertise of the avatar influenced perceptions of credibility, and credibility mediated the influence of the avatar's expertise on persuasion.

The results raise questions about persuasion in a Web-based environment. People are often aware of attempts at persuasion and have developed several strategies for resisting persuasion (Jacks and Cameron 2003; Knowles and Linn 2004). First, when people identify an attempt at persuasion, they can discount the message (Gruder et al. 1978) or derogate the source (Tannenbaum, Macauley, and Norris 1966; Zuwerink and Devine 1996). Second, people can use selective attention to attitude-congruent information or selective avoidance of attitude-incongruent information (Albarracín and Mitchell 2004). Third, people can have negative affective responses (e.g., irritation, anger) when exposed to a persuasive message, which leads to the rejection of the message (Zuwerink and Devine 1996). Fourth, people can actively look for a fault in the persuasive message and generate counterarguments (Brock 1967; Killeya and Johnson 1998). Finally, people can bolster their own attitude by selectively generating or recalling information in support of their own viewpoint (Lewan and Stotland 1961; Lydon, Zanna, and Ross 1988). These latter two strategies are likely to produce enduring resistance. However, they require cognitive resources and motivation.

Our two studies show that avatars can be effective instruments of persuasion, but further research needs to be conducted to answer the following questions: (1) To what extent do avatars have the potential to evade the common strategies consumers use to resist persuasive messages? and (2) What factors support consumer acceptance of avatars as helpful sales agents? In particular, two classes of factors should be taken into consideration. First, there are many options for the design of virtual characters in Web-based environment in terms of function, graphic design, and technical organization. It is likely that the persuasive power of avatars varies significantly, depending on whether avatars deliver a testimonial or act as representatives of the company, whether avatars are designed as more or less lifelike, and whether avatars provide one-way or reciprocal conversation. Second, recipient characteristics influence persuasion. That is, different people have different motivations and abilities to process information and to resist persuasion.

Finally, five research issues need discussion. First, in our procedure, we did not collect sales response data. Ideally, an investigation into avatar effectiveness would create a Web site with an avatar and a Web site without an avatar and show that sales are higher on an avatar-accompanied Web site. Second, we generated the results using a single product category in which a consultation was beneficial. We expect that the influence of an avatar will decline as the purchase process becomes easier, the product becomes simpler, and the buyer's knowledge of the product increases. Third, it could be argued that the avatar effects were novelty effects because few respondents had experienced avatar-accompanied Web sites. This may be true, but it is also the case that avatars varied in their effectiveness as a result of different levels of involvement. This interaction is inconsistent with a simple novelty effect. Fourth, we expect that avatars are effective because they enhance perceptions of reciprocal communication. Although our avatars engaged shoppers by taking turns conversationally with them, the exchange fell short of a truly reciprocal communication in which participants respond to the idiosyncratic conversation of their partner. Fifth, it could be argued that the results show no incremental conceptual insight relative to the existing literature on sales force effectiveness. However, if this premise is accepted, it must also be acknowledged that avatars may have many of the same benefits as human sales agents. This was the goal of the article.

Appendix A
Questionnaire Items

Product Involvement (α = .90)

- "For me, customized leisure shoes are (important) (fun) (exciting) (relevant)."

Avatar Attractiveness (α = .90)

- "In my opinion, the virtual consultant is (attractive) (beautiful) (good looking)."

Avatar Expertise (α = .93)

- "In my opinion, the virtual consultant is (trained) (experienced) (knowledgeable)."

Entertainment Value of the Web Site (α = .93)

In your opinion, to what extent do the following statements about the content and design of the Web-site apply to you?

- "I find the Web-site entertaining."
- "I like the Web-site."
- "The Web-site is fun to use."

Information Value of the Web Site (α = .85)

In your opinion, to what extent do the following statements about the content and design of the Web-site apply to you?

- "The information offered is useful."
- "The information offered is understandable."
- "The information offered is sufficient."

Avatar Likeability (α = .89)

- "In my opinion, the virtual consultant is (likeable) (friendly) (agreeable)."

Avatar Credibility (α = .94)

- "In my opinion, the virtual consultant is (sincere) (competent) (credible)."

Satisfaction with the Retailer (α = .94)

"Considering everything that I have learned so far about the company and its products, my opinion is that"

- "The company fulfills my needs."
- "It would be advantageous for me to buy products from this company."
- "I am delighted about the company and its products."
- "I am satisfied with the company."

Attitude Toward the Product (α = .94)

- "I find the company's customized leisure shoes (useful) (agreeable) (favorable) (good)."

Intention to Buy (α = .95)

- "I can imagine buying a pair of customized leisure shoes from this company."
- "The next time I buy a pair of leisure shoes, I will take this company into consideration and have them make me an offer."
- "I am very interested in buying a pair of customized leisure shoes from this company."

Appendix B[13]

Low-Expertise Consultation

Introduction: Hi, my name is Kim. / I'm a new shoe salesperson and I am not yet very familiar with this merchandise. I can, however, tell you how to learn more about our awesome shoes. / Our shoes can be made exactly as you want them. The shoes will have just the right style and they will be comfortable. Your feet will feel wonderful in them. Find out more about the advantages of custom-made shoes.

Page 1: A person's appearance makes a strong first impression. Therefore, the appearance of a shoe is often the most important factor in a shoe purchase. Unfortunately, stylish shoes are often uncomfortable. There's nothing worse than sore feet at the end of a long day. We use a certain technique for custom-fit production so you can get shoes that feel great and look the way you want them to. / Are you looking for a women's or men's shoe?

Page 2: You can design your shoe by using the mouse to click on the right. Here is a quick tip. If there is something that you are unsure about, click on the "i" and an explanation will be provided. / To begin with, there is a shoe in white material that is already pretty well equipped. I recommend the Creative Design option for you, even though it can be a bit complicated if one is not a professional in using this design procedure. There are, after all, over 50 colors from which you can choose. Still, it's fun and the result is sure to be a one-of-a-kind shoe. / You can use the Creative Design after you have selected all of the other options for your shoe.

Page 3: The inside of the shoe can be made either out of a basic fabric or out of a waterproof material. / The soles are important for the comfort of the shoe. Therefore, you can now choose some features to make the soles of your shoes even better. This is important for people that spend a lot of time on their feet. / After you made your selections, I can calculate a price for you.

Page 4: Thanks to our online store, I can offer your desired shoes at a very inexpensive price. / You have now completed the shoe consultation. We would like to ask you some questions about this experience.

High-Expertise Consultation

Introduction: Hello, my name is Dr. Anne Schneider. / I have a doctorate in sports science. I have been an orthopedist and podiatrist for over 15 years. I would like to offer you comprehensive advice about our innovative leisure shoes. / These shoes can be made according to your needs.

---

[13] A "/" represents a page break.

Our shoes provide optimal support to your feet and prevent back strain. Find out more about the advantages of custom-made shoes.

Page 1: The fit of a shoe is a very important factor when buying shoes. Every day, your feet are exposed to enormous strain, which all-too-often causes discomfort. Orthopedically defective shoes can cause idiopathic scoliosis as well as degenerative symptoms like arthritis. / Our shoes are custom-made using innovative technologies. They have been designed to provide optimal support to your foot structure and thus prevent discomfort. The design is also based on your preferences.

Page 2: You can design your shoe by using the mouse to click on the right. All options are explained by clicking on the info button "i". / As a base model, our experts have developed a textile shoe in white with a high-quality configuration. I recommend the 3-D Fit option for you. This option guarantees that the shoe will fit optimally and offer maximum wearing comfort. These are essential prerequisites for healthy and relaxed walking.

Page 3: The material inside of the shoe can be made of a basic fabric or a special, water-resistant, high-tech membrane. / Because the quality of the soles determines the comfort and walking characteristics of the shoes, you can choose to upgrade to our high-quality standard sole. Silicon absorption, for example, reduces the incidence of back pain up to 20%. / After you made your selections, I will gladly calculate an exact price for your shoe.

Page 4: Thanks to our online customization and a revolutionary production process, I can offer you customized shoes at a very attractive price. / You have now completed the shoe consultation. We would like to ask you some questions about this experience.

REFERENCES

Albarracín, Dolores and Amy L. Mitchell (2004), "The Role of Defensive Confidence in Preference for Proattitudinal Information: How Believing That One Is Strong Can Sometimes Be a Defensive Weakness," *Personality and Social Psychology Bulletin*, 30 (December), 1565–84.

Allen, Mike, Raymond W. Preiss, Barbara Mae Gayle, and Nancy A. Burrell (2002), *Interpersonal Communication Research: Advances Through Meta-Analysis*. Mahwah, NJ· Lawrence Erlbaum Associates.

Bahorsky Russ, Jeffrey Graber, and Steve Mason (1988), *Official Internet Dictionary: A Comprehensive Reference for Professionals Reviews* Portland: ABS Consulting.

Barlow, Alexis K.J., Noreen Q. Siddiqui, and Mike Mannion (2004), "Development in Information and Communication Technologies for Retail Marketing Channels," *International Journal of Retail and Distribution Management*, 32 (March), 157–63.

Baron, Reuben M. and David A. Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (December), 1173–82.

Brehm, Sharon S., Saul M. Kassin, and Steven Fein (2005), *Social Psychology*. Boston: Houghton Mifflin.

Brock, Timothy C. (1967), "Communication Discrepancy and Intent to Persuade as Determinants of Counterargument Production," *Journal of Experimental Social Psychology*, 3 (3), 296–309.

Brown, Steven and Douglas Stayman (1992), "Antecedents and Consequences of Attitude Toward the Ad: A Meta-Analysis," *Journal of Consumer Research*, 19 (June), 4–51.

Bull, Ray and Nichola Rumsey (1988), *The Social Psychology of Facial Appearance*. New York: Springer.

Caballero, Marjorie J. and William M. Pride (1984), "Selected Effects of Salesperson Sex and Attractiveness in Direct Mail Advertisements," *Journal of Marketing*, 48 (Winter), 94–100.

Chaiken, Shelly (1979), "Communicator Physical Attractiveness and Persuasion," *Journal of Personality and Social Psychology*, 37 (August), 1387–97.

Childers, Terry L., Christopher L. Carr, Joann Peck, and Stephen Carson (2001), "Hedonic and Utilitarian Motivations for Online Retail Shopping Behavior," *Journal of Retailing*, 77 (Winter), 511–35.

Cooper, Joel and Robert T. Croyle (1984), "Attitudes and Attitude Change," *Annual Review of Psychology*, 35, 395–426.

Dion, Kenneth K., Ellen Berscheid, and Elaine Walster (1972), "What Is Beautiful Is Good," *Journal of Personality and Social Psychology*, 24 (December), 285–90.

Dykema, Evie B. (2000), "Online Retail's Ripple Effect," Research Report, Forrester Research Inc., Cambridge, MA (September 2000).

Eighmey, John and Lola McCord (1998), "Adding Value in the Information Age: Uses and Gratifications of Sites on the World Wide Web," *Journal of Business Research*, 41 (March), 187–94.

Freiden, Jon B. (1984), "Advertising Spokesperson Effects: An Examination of Endorser Type and Gender on Two Audiences," *Journal of Advertising Research*, 24 (October–November), 33–41.

Friedman, Hershey H. and Linda Friedman (1979), "Endorser Effectiveness by Product Type," *Journal of Advertising Research*, 19 (October), 63–71.

G&J Electronic Media Services (2001), *7th Wave of the GfK-Online-Monitor*. Hamburg: GfK Press

Gruder, Charles L., Thomas D. Cook, Karen M. Hennigan, Brian R. Flay, Cynthia Alessis, and Jerome Halamaj (1978), "Empirical Tests of the Absolute Sleeper Effect Predicted from the Discounting Cue Hypothesis," *Journal of Personality and Social Psychology*, 36 (October), 1061–1074.

Halfhill, Tom R. (1996), "Agents and Avatars," *Byte*, 21 (February), 69–72.

Horai, Joan M., Nicolas Naccari, and Elliott Fatoullah (1974), "The Effect of Expertise and Physical Attractiveness upon Opinion Agreement and Liking," *Sociometry*, 37 (December), 601–606.

Horton, Donald and Richard R. Wohl (1956), "Mass Communication and Para-Social Interaction," *Psychiatry*, 19 (August), 215–29.

Hovland, Carl I., Irving L. Janis, and Harold H. Kelley (1953), *Communication and Persuasion* New Haven, CT: Yale University Press.

Jacks, Julia Zuwerink and Kimberly A. Cameron (2003), "Strategies for Resisting Persuasion," *Basic and Applied Social Psychology*, 25 (June), 145–61.

Joseph, W. Benoy (1982), "The Credibility of Physically Attractive Communicators: A Review," *Journal of Advertising*, 11 (3), 15–24.

Katz, Elihu and Paul F. Lazarsfeld (1955), *Personal Influence*. New York: The Free Press.

Kelman, Herbert C. (1961), "Processes of Opinion Change," *Public Opinion Quarterly*, 25 (Spring), 57–78.

Killeya, Ley A. and Blair T. Johnson (1998), "Experimental Induction of Biased Systematic Processing: The Directed-Thought Technique," *Personality & Social Psychology Bulletin*, 24 (January), 17–33.

Knowles, Eric S. and Jay A. Linn (2004), *Resistance and Persuasion*. Mahwah, NJ: Lawrence Erlbaum Associates.

Korgaonkar, Pradeep K. and Lori D. Wolin (1999), "A Multivariate Analysis of Web Usage," *Journal of Advertising Research*, 3 (March–April), 53–68.

Lewan, Paul C. and Ezra Stotland (1961), "The Effects of Prior Information on Susceptibility to an Emotional Appeal," *Journal of Abnormal and Social Psychology*, 62 (March), 450–53.

Loos, Peter (2003), "Avatar," in *Lexikon Electronic Business*, Thomas Schildhauer, ed. Munich: Oldenbourg, 16–19.

Lydon, John, Mark P. Zanna, and Michael Ross (1988), "Bolstering Attitudes by Autobiographical Recall: Attitude Persistence and Selective Memory," *Personality and Social Psychology Bulletin*, 14 (March), 78–86.

MacCallum, Robert C., Shaobo Zhang, Kristopher J. Preacher, and Derek D. Rucker (2002), "On the Practice of Dichotomization of Quantitative Variables," *Psychological Methods*, 7 (March), 19–40.

MacKenzie, Scott B. and Richard J. Lutz (1989), "An Empirical Examination of the Structural Antecedents of Attitude Toward the Ad in an Advertising Pretesting Context," *Journal of Marketing Research*, 53 (November), 48–65.

———, ———, and George E. Belch (1986), "The Role of Attitude Toward the Ad as a Mediator of Advertising Effectiveness: A Test of Competing Explanations," *Journal of Marketing Research*, 23 (May), 130–43.

Maddux, James E. and Ronald W. Rogers (1980), "Effects of Source Expertness, Physical Attractiveness, and Supporting Arguments on Persuasion: A Case of Brains over Beauty," *Journal of Personality and Social Psychology*, 39 (August), 235–44.

McGuire, William J. (1985), "Attitudes and Attitude Change," in *The Handbook of Social Psychology*, Vol. 2, Lindzey Gardner and Elliot Aronson, eds. New York: Knopf, 233–346.

Mills, James and John Harvey (1972), "Opinion Change as a Function of When Information About the Communicator Is Received and Whether He Is Attractive or Expert," *Journal of Personality and Social Psychology*, 21 (January), 52–55.

Moon, Youngme (2000), "Intimate Exchanges: Using Computers to Elicit Self-Disclosure from Consumers," *Journal of Consumer Research*, 26 (March), 323–39.

——— (2003), "Don't Blame the Computer: When Self-Disclosure Moderates the Self-Serving Bias," *Journal of Consumer Psychology*, 13 (1–2), 125–37.

Morkes, John, Hadyn K. Kernal, and Clifford I. Nass (1999), "Effects of Humor in Task-Oriented Human-Computer Interaction and Computer-Mediated Communication: A Direct Test of SRCT Theory," *Human-Computer Interaction*, 14 (4b), 395–435.

Mummert & Partner (2001), "Virtuelle Assistenten: Geheimwaffe gegen Milliardenverluste im Internet," press release, (March 27), (accessed April 22, 2005), [available at www.news4press.net/presse/985726200694.php3].

Nass, Clifford I., Matthew Lombard, Lisa Henriksen, and Jonathan S. Steur (1995a), "Anthropocentrism and Computers," *Behaviour and Information Technology*, 14 (4), 229–38.

———, Youngme Moon, B.J Fogg, Byron Reeves, and Chris Dryer (1995b), "Can Computers Be Human Personalities?" *International Journal of Human-Computer Studies*, 43 (August), 223–39.

——— and Jonathan Steur (1993), "Voices, Boxes and Sources of Messages: Computers and Social Actors," *Human Communication Research*, 19 (June), 504–527.

Nielsen//NetRatings (2005), "Online Retail Report Card: Highest Conversion Rates Amongst Non-Brick-and-Mortar Retailers, According to Nielsen//NetRattings Megaview Online Retail," press release, (April 7), (accessed April 22, 2005), [available at http://www.netratings.com/pr/pr_050407.pdf].

Petty, Richard E., John T. Cacioppo, and Rachel Goldman (1981), "Personal Involvement as a Determinant of Argument-Based Persuasion," *Journal of Personality and Social Psychology*, 41 (November), 847–55.

———, ———, and Martin Heesacker (1981), "Effects of Rhetorical Questions on Persuasion: A Cognitive Response Analysis," *Journal of Personality and Social Psychology*, 40 (March), 432–40.

Phillips, Fred, Andrew Donoho, William W. Mayberry, Walter Keep, John M. McCann, Karen Shapiro, and David Smith (1997), "Electronically Connecting Retailers and Customers: Interim Summary of Expert Roundtable," in *Electronic Marketing and the Consumer*, Robert A. Peterson, ed. Thousand Oaks, CA: Sage Publications, 101–122.

Redmond, William H. (2002), "The Potential Impact of Artificial Shopping Agents in E-Commerce Markets," *Journal of Interactive Marketing*, 16 (December), 56–66.

Reeves, Byron and Clifford I. Nass (1996), *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge, MA: Cambridge University Press.

Reingen, Peter H. and Jerome B. Kernan (1993), "Social Perception and Interpersonal Influence: Some Consequences of the Physical Attractiveness Stereotype in a Personal Selling Setting," *Journal of Consumer Psychology*, 2 (1), 25–39.

Reynolds, Kristy E. and Sharon E. Beatty (1999), "Customer Benefits and Company Consequences of Customer-Salesperson Relationships in Retailing," *Journal of Retailing*, 75 (Spring), 11–32.

Rubin, Alan M., Elizabeth Perse, and Robert A. Powell (1985), "Loneliness, Parasocial Interaction, and Local Television News Viewing," *Human Communication Research*, 12 (Winter), 155–80.

Rubin, Vicki, Carol Mager, and Herschey H. Friedman (1982), "Company President Versus Spokesperson in Television Commercials," *Journal of Advertising Research*, 22 (August–September), 31–33.

Soldow, Gary F. and Gloria Penn Thomas (1984), "Relational Communication: Form Versus Content in the Sales Interaction," *Journal of Marketing*, 48 (Winter), 84–93.

Stäudel, Thea (1987), *Problemlösen, Emotionen, Kompetenz: Die Überprüfung eines Integrativen Konstrukts*. Regensburg: Roderer.

Stephens, Nancy and William T. Faranda (1993), "Using Employees as Advertising Spokespersons," *Journal of Services Marketing*, 7 (2), 36–46.

Swaminathan, Vanitha, Elzbieta Lepkowska-White, and Bharat P. Rao (1999), "Browsers or Buyers in Cyberspace? An Investigation of Factors Influencing Electronic Exchange," *Journal of Computer-Mediated Communication*, 5 (December), (accessed April 22, 2005), [available at http://jcmc.indiana.edu/vol5/issue2/swaminathan.htm].

Tannenbaum, Percy H., Jacqueline R. Macauley, and Eleanor L. Norris (1966), "Principle of Congruity and Reduction of Persuasion," *Journal of Personality and Social Psychology*, 3 (2), 233–38.

Trogemann, Georg (2003), "Mit Hand und Fuß: Die Bedeutung der nonverbalen Kommunikation für die Emotionalisierung von Dialogführungssystemen," in *Avatare-Digitale Sprecher für Business und Marketing*, Christian Lindner, ed. Berlin: Springer, 267–90.

U.S. Census Bureau (2005), "Retail 4Q, 2004 E-Commerce Report," *U.S. Census Bureau*, (accessed April 22, 2005), [available at http://www.census.gov/mrts/www/ecomm.html].

Webster, Frederick E. (1968), "Interpersonal Communication and Salesman Effectiveness," *Journal of Marketing*, 32 (July), 7–13.

Weitz, Barton A., Harish Sujan, and Mita Sujan (1986), "Knowledge, Motivation, and Adaptive Behavior: A Framework for Improving Selling Effectiveness," *Journal of Marketing*, 50 (October), 174–91.

Weizenbaum, Joseph (1976), *Computer Power and Human Reason. From Judgment to Calculation.* New York: Penguin.

Williams, Ederyn (1977), "Experimental Comparisons of Face-to-Face and Mediated Communications," *Psychological Bulletin*, 84 (September), 963–76.

Wood, Wendy (2000), "Attitude Change: Persuasion and Social Influence," *Annual Review of Psychology*, 51, 539–70.

Zuwerink, Julia R. and Patricia G. Devine (1996), "Attitude Importance and Resistance to Persuasion: It's Not Just the Thought That Counts," *Journal of Personality and Social Psychology*, 70 (May), 931–44.

Gary F. Gebhardt, Gregory S. Carpenter, & John F. Sherry Jr.

Creating a Market Orientation: A Longitudinal, Multifirm, Grounded Analysis of Cultural Transformation

Market orientation is a foundation of marketing and is increasingly important in other fields, such as strategic management. Research in marketing has identified the characteristics of market-oriented organizations. However, how organizations change to become more market oriented has received less attention. In this article, the authors conduct an in-depth, longitudinal, multifirm investigation of firms that have successfully created a market orientation. Grounded by this in-depth understanding, they develop a theoretical model to explain how firms create a market orientation. The model identifies four path-dependent stages of change. In contrast to current conceptualizations, the authors find that creating a market orientation requires dramatic changes to an organization's culture and the creation of organizationally shared market understandings. The findings offer new insights into how organizations develop a greater market orientation, organizational change, and the nature of market orientation, including the role of intraorganizational power and organizational learning in creating and sustaining a market orientation.

The concept of a market orientation has become increasingly important to the study and practice of management. It is the central concept of marketing (e.g., Kotler 2000), it has become increasingly relevant to scholars in other fields such as management (e.g., Besanko, Dranove, and Shanley 2000), and it is recognized as significant for organizational success (e.g., Collins and Porras 1994). For example, Intel's chief executive officer (CEO) recently announced plans to transform the company so that "every idea and technical solution should be focused on meeting customers' needs from the outset" (Edwards 2005, p. 35). Organizations such as Intel presumably hope to reap the rewards that research has shown to be associated with a greater market orientation (e.g., Narver and Slater 1990).

Despite the great and growing interest in this marketing concept, research on how organizations become more market oriented is surprisingly limited. Much research on market orientation has focused on developing measures of a firm's orientation and identifying antecedents and consequences of a greater market orientation (e.g., Homburg and Pflesser 2000; Kirca, Jayachandran, and Bearden 2005; Kohli and Jaworski 1990; Kohli, Jaworski, and Kumar 1993). These studies provide important insights into the dimensions of market orientation and its implications, but they are cross-sectional analyses and therefore cannot offer insights into the dynamics of organizational change (Jaworski and Kohli 1993). Using a prescriptive approach, Day (1999) develops a more expansive model of organizational change to a greater market orientation. Consistent with previous work in marketing, the model focuses on creating a more market-oriented organization through formal management actions, such as changing incentives and organizational structures.

Research on organizational change suggests that though formal actions are an important avenue of change, the change process is much more complex. Organizations change as a result of many forces, including political struggles among factions within the organization, evolutionary change through organizational experimentation and learning, social pressures from outside the organization, and the changes that naturally occur as an organization matures through a life cycle (Van de Ven and Poole 1995; Zald and Berger 1978). For example, disempowered members of an organization can create an insurgency that erodes the power of the leadership until it collapses. Similarly, managers can mobilize external pressure to force organizational change. Such mechanisms play documented roles in corporate, governmental, and institutional change (Zald and Berger 1978). How or, indeed, whether these mechanisms operate and in

---

Gary F. Gebhardt is Assistant Professor of Marketing, College of Business, University of South Florida (e-mail: ggebhard@coba.usf.edu). Gregory S. Carpenter is James Farley/Booz Allen Hamilton Professor of Marketing Strategy and Director of the Center for Market Leadership, Kellogg School of Management, Northwestern University (e-mail: g-carpenter@northwestern.edu). John F. Sherry Jr. is Raymond W. & Kenneth G. Herrick Professor of Marketing and Marketing Department Chair, Mendoza College of Business, University of Notre Dame (e-mail: jsherry@nd.edu). This article is based on Gary F. Gebhardt's dissertation, cochaired by Gregory S. Carpenter and John F. Sherry Jr. The authors thank Ed Zajac, Paul Hirsch, Mike Barnett, Anand Kumar, Rajiv Dant, Hayagreeva Rao, Jim Stock, François Carrillat, Cristian Chelariu, Vicki Panzer, and Ryon Lancaster for helpful insights and comments. The authors also thank the Institute for the Study of Business Markets and the Marketing Science Institute for financial and other research support. Most of all, the authors thank the individuals from the seven participating organizations who contributed their time and effort to this research.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

what combination or sequence remains largely unexplored in the context of creating a greater market orientation.

In this article, we explore the process through which organizations change to adopt a greater market orientation by studying a group of firms over time. Each of these firms was engaged in a process to create a greater market orientation, though different firms were at different stages of the process. To understand the changes that each was undergoing and to deduce a common process, we studied each firm using the longitudinal-processual method of in-depth, qualitative examination, including ethnographic observation, oral histories, and analysis of historical documents. To do so, we relied on methods that are common in consumer behavior, anthropology, sociology, and organizational research (e.g., Glaser and Strauss 1967; Pettigrew 1990; Sherry and Kozinets 2001; Thompson, Locander, and Pollio 1989). On the basis of this analysis, we identified a common, longitudinal process that leads to a greater market orientation. We identified the stages of this process, the obstacles impeding progress from one stage to the next, and the characteristics of firms that successfully navigate the process.

Our analysis reveals a four-stage process of organizational change: (1) initiation, (2) reconstitution, (3) institutionalization, and (4) maintenance. We describe each of the four stages (and their linkages and implications) in more nuanced detail subsequently, but the process can be briefly summarized as follows: Members of the organization recognize a threat to the organization. Threats take many forms, but they are typically not widely recognized by the organization. To address the threat, a group of empowered managers creates a coalition to plot the change process. The managers plan the complete transformation of the organization, mobilize the larger organization, and create a cultural shift through a process of reconnecting organization members with customers. Through this process, they build common experiences and perspectives and, at the same time, build a growing consensus for more formal organizational changes. These formal changes follow informal ones and are used to sustain the new orientation of the organization.

The process we identify has important implications for our understanding of creating a more market-oriented organization and for the marketing concept in general. Previous work on market orientation identifies three factors as antecedents of a greater market orientation: top management focus, interdepartmental cooperation, and reward systems (Kirca, Jayachandran, and Bearden 2005). Each of these factors is believed to operate independently to enhance market orientation. Our analysis finds these to be among a larger set of factors that operate to create organizational change. Moreover, we find that these factors are embedded in a larger, richer sequence of change processes. Our analysis describes this fuller range of factors and the sequential process in which they are embedded. For example, the change process we describe depends on the formation of an elite group of top managers who are determined to change the organization in its most fundamental ways, including the values of the organizational culture. We find that reward systems are an important element of the process but only after top management has created greater support for change throughout the organization. We describe each stage in detail, examine the sequence of changes, and discuss the implications of the richness revealed in the change process.

Our results offer new insights into the process of organizational change. Scholars in organizational behavior, sociology, and other fields have proposed different mechanisms and modes through which organizations change, including evolutionary forces and political struggle (Van de Ven and Poole 1995; Zald and Berger 1978). However, which mechanisms or modes of change may or may not play a role in creating a market-oriented organization and how they do so remains unresolved. Our analysis describes which change mechanisms, which modes of change, and which sequences are part of creating a more market-oriented organization. For example, we discuss how the process of formal change in organizational structure and rewards follows a process of mass mobilization led by senior management. Although the mechanisms we identify are clearly related to those previously identified, they differ in important ways in the specific context we examine. We discuss these differences and their implications.

Our results offer new insights into the nature of a market orientation. Marketing scholars have debated whether market orientation is a set of behaviors or a culture (e.g., Deshpandé and Farley 1998; Homburg and Pflesser 2000; Narver and Slater 1998). Our analysis shows that market orientation rests fundamentally on cultural values. We identify six key values that are shared by the market-oriented organizations in our analysis. These values, which are shared throughout the organization, become central to the change process. Indeed, an early objective of people who transform organizations is to establish new values and norms for behavior. These values are central to a market orientation. Unlike previous studies, our analysis shows that the adoption of these values leads to a redistribution of power through the organization, a change in the nature of organizational learning, and a change in the organization's capabilities. Our analysis illustrates that the distribution of intraorganizational power and organizational learning are essential, overlooked characteristics of market-oriented firms. We explore the role of each of these elements in the creation of a market-oriented culture.

Methodology and Research Design

To discover how firms create a market orientation, we observed and analyzed firms seeking to create a greater market orientation. We conducted our research using the longitudinal-processual method (Pettigrew 1979, 1990). This approach combines in-depth, qualitative data collection techniques (i.e., ethnographic observation, depth interviews [in particular, oral histories], and historical documents) with comparative analysis to develop a process model of change. The longitudinal-processual method compares each organization with itself at different stages of a chronological process and compares longitudinal progressions across organizations to develop insights into a common change process. Organizational behavior research has used this approach to study multiyear processes, such as internal corporate venturing (Burgelman 1983) and the development of organizational cultures (Pettigrew 1979). Rather than specify possible change mechanisms a priori, as previous studies have done, our approach is distinctly inductive. Grounded by an in-depth understanding of how a sample of firms have changed over time, we develop a model for how firms become market oriented. On the basis of this analysis, we produce a generalized understanding that can be empirically verified in subsequent research (Deshpandé 1983; Glaser and Strauss 1967).

The ethnographic investigation of behavior in complex organizations has a long and discontinuous history in the United States, dating back to the late 1920s. Periodic calls for the revival and sustaining of the tradition have arisen from within the social sciences (Whyte 1978) and business (Bonoma 1985) disciplines and have fueled recent inquiry both from sociologists and anthropologists who are interested in business dynamics and from marketing scholars who are intrigued by a nonpositivist understanding of managerial motivation. Epistemological analyses, methodological tutorials, and exemplary studies (Arnould and Thompson 2005; Arnould and Wallendorf 1994; Belk, Sherry, and Wallendorf 1988; Deshpandé and Webster 1989; Sherry 1995, 1998; Sherry and Kozinets 2001) have established ethnography as part of marketing's canonical toolkit. Our investigation is part of this current wave of qualitative research into marketing.

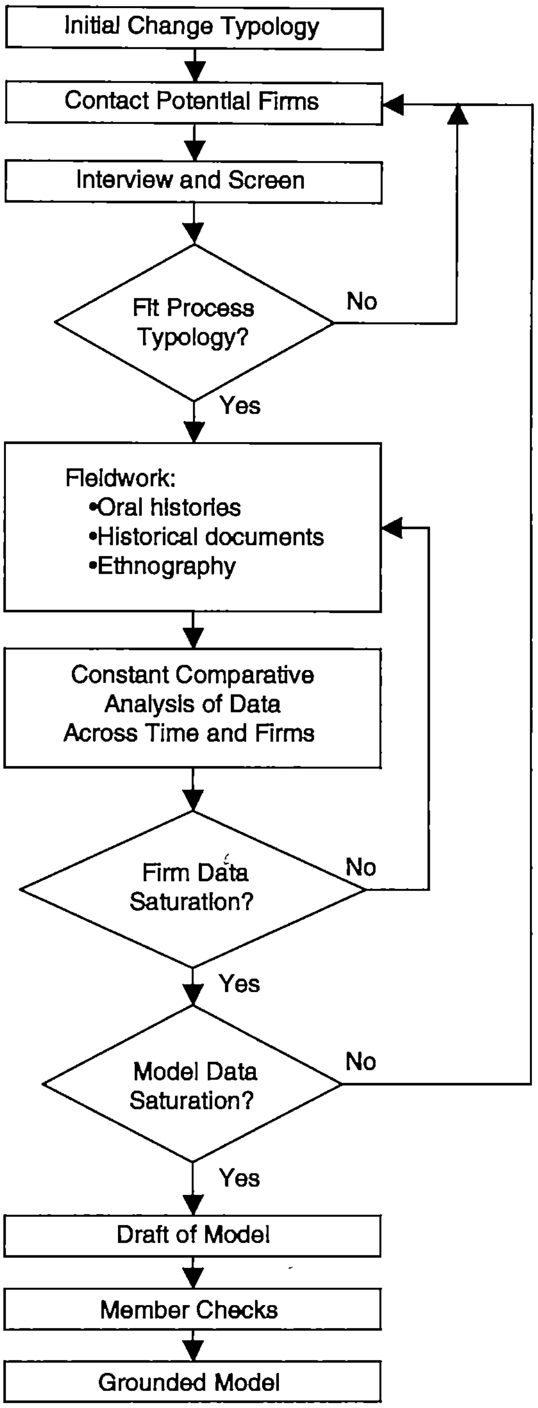
Research Process

Figure 1 provides an overview of the way we implemented the longitudinal-processual method for this research. Our first step was to develop a typology of organizations that might offer the greatest insight into the process of creating a market orientation. Knowing that four to six firms would be the practical maximum we could work with given the longitudinal-processual method, we followed a theoretical sampling approach (Eisenhardt 1989; Strauss and Corbin 1998). Assuming a simple longitudinal model of change, we considered firms at three stages: beginning the change process, currently changing, and having achieved a market orientation. We attempted to recruit two firms in each stage. We identified potential firms on the basis of academic and practitioner recommendations, reviews of the business press, and company documents. From that larger pool of firms, we generated a pool of willing candidates.

MediaCo and EquipmentCo (pseudonyms) were firms beginning the change process. Both firms recently launched initiatives with the explicit goal of creating a market orientation. EquipmentCo is a capital equipment manufacturer selling to business markets, and MediaCo develops communication products and services for consumers on behalf of corporate and not-for-profit customers.

We screened firms that were in the process of changing or had completed a change to a market orientation through key manager interviews. During this screening, we ascertained whether the firms experienced increases in market orientation over a period of time relative to Kohli and Jaworski's (1990, p. 6) definition of market orientation as the "organizationwide *generation* of market intelligence pertaining to current and future customer needs, *dissemination* of the intelligence across departments, and organizationwide *responsiveness* to it." These interviews helped mitigate demand characteristics associated with questionnaires and prevented contamination of prospective field sites.

FIGURE 1
Research Process



On the basis of these conversations, we recruited two firms that were in the process of change, BenefitsInc (a pseudonym) and Motorola Personal Communications Sector (PCS), and two firms that had recently completed a transformation, Alberto-Culver and Marshfield Door-

Systems. BenefitsInc is an employee benefit company that was created in the 1980s through the merger of two smaller organizations. Within Motorola, Motorola PCS develops and manufactures wireless communication devices, which are typically sold to wireless carriers that resell them to consumers. Alberto-Culver is a multinational manufacturer and marketer of consumer and professional hair care and beauty products, such as Alberto VO5. Our research focused on the experiences of Alberto-Culver North America. Marshfield DoorSystems is a privately held commercial architectural door manufacturer in Marshfield, Wis., which was a Weyerhaeuser subsidiary when it began its change efforts.

Data collection included gathering oral histories, reviewing historical documents, and engaging in ethnographic observation.[1] We obtained oral histories by interviewing a cross-section of employees, both vertically (CEOs and shop floor workers) and horizontally (employees from marketing, finance, engineering, and so forth). The majority of interviewees joined their organization before it began transforming to a market orientation. Oral history interviews lasted from 40 minutes to three hours and were audiotaped. We conducted more than 70 formal interviews during ten months of fieldwork, representing more than 120 hours of audio recordings.

We began interviews in an exploratory manner, focusing on each individual's phenomenological interpretation of the firm's transformation and current state (e.g., Glaser and Strauss 1967; Thompson, Locander, and Pollio 1989). When an informant shared an interpretation of the firm and how it had changed, we clarified ambiguities and summarized what we heard, providing an opportunity for informants to correct or expound on anything we might have misunderstood. We then inquired about factual (e.g., a specific event) and interpretive (e.g., what the event meant) differences among informants within an organization. This provided informants with an opportunity to address events they had forgotten to mention, explain why different people within the same organization interpreted an event differently, and help us make sense of the changes occurring within each organization. Having obtained informants' phenomenological interpretations of change, we concluded the interviews by explaining the nature of the project in more detail, including what we were finding across firms. By sharing this information, we sought to ensure that we elicited any information or interpretation that might be relevant to understanding the transformation to a market orientation. Leveraging the discovery-oriented approach with the vast literature on how other types of organizational change occur, we moved from a phenomenological to a directed inquiry approach, ensuring that we explicitly considered prior suggestions of how firms might change to become more market oriented (Eisenhardt 1989; Strauss and Corbin 1998).

We reviewed hundreds of historical documents across participating firms. External documents included annual reports, company press releases, industry publications, and articles from both national and local media. Internal company documents included memos, strategic plans, training materials, and marketing research. Historical documents provided context regarding activities and attitudes in firms before and during their transformations, as well as confirmed and elaborated on oral histories. In addition, historical internal documents provided deeper insight into the actual process of change at each firm.

We conducted ethnographic fieldwork, observing and interacting with firms' employees, customers, channels, suppliers, and consultants. We spent more than 40 days in the field, sitting in senior executive, marketing strategy, sales review, and consultant meetings and sharing breakfasts, coffees, and lunches with informants in company cafeterias, boardrooms, and local eateries. We spent time observing and speaking with employees in company work and break areas. We toured facilities and initiated conversations with whomever we met. We even worked as Harley-Davidson employees at Daytona Bike Week, where we also joined informants for dinners and drinks. Participant observation in the field gave insight into the lived experience of employees engaged in everyday activities.

Data collection and analysis occurred over ten months. We managed interview transcriptions, field notes, and historical documents electronically, using QSR International's NVivo software, which allowed us to code and analyze a vast amount of data across participating firms. Following standard practice for qualitative inquiry, we coded the data iteratively, seeking to find common themes across firms. We refined those themes on the basis of further fieldwork and ongoing discussion (Eisenhardt 1989; Strauss and Corbin 1998). For the initial interviews, field notes, and historical documents, we selectively coded for anticipated change mechanisms (e.g., incentives) and openly coded other passages that appeared to be important to the change process on the basis of informant recollections or our own observations. As we added information from additional sources or incremental firms, we continued to refine our understanding of common processes across firms, which we captured by changes to our coding scheme (Eisenhardt 1989; Strauss and Corbin 1998). Using this approach, we developed an understanding of change at each firm. Simultaneously, we developed a common model of change to a market orientation by comparing and contrasting our findings across firms in a discovery-based approach to theory development (Glaser and Strauss 1967). When we understood how a given firm had changed and our comparisons raised no further questions, we agreed that our understanding had reached saturation, and we concluded fieldwork with that firm (Glaser and Strauss 1967; Strauss and Corbin 1998).

An advantage of the grounded theory methodology and a theoretical sampling paradigm is the flexibility to alter research design on the basis of an evolving understanding of the change process (Glaser and Strauss 1967; Strauss and Corbin 1998). For example, after a period of fieldwork, we identified the fourth stage of change, unanticipated in our original design: maintaining a market orientation. On the basis of this discovery, we recruited Harley-Davidson as a seventh participating firm. Harley-Davidson, a $4 billion company that provides motorcycles, branded products, and services, underwent tremendous organizational change during the 1980s to become market oriented.

---

[1] The first author conducted all data collection and analysis under the direction of his dissertation cochairs. We employ the pronoun "we" throughout the article to minimize distraction.

During our fieldwork, it also became evident that BenefitsInc, EquipmentCo, and MediaCo were not generating, disseminating, or responding to market intelligence consistent with a market orientation (Kohli and Jaworski 1990), and they had not made any discernable progress toward creating a market orientation. Change efforts at BenefitsInc were isolated to branding and brand management activities. Although change efforts at EquipmentCo and MediaCo began more than a year before our fieldwork, these efforts did not seem to be having much impact at either firm. Progress appeared to be impeded by a lack of sufficient political power to implement change and a lack of understanding throughout each organization regarding the nature of the desired change. Nonetheless, an understanding of the change efforts provides a valuable contrast with firms that succeeded in creating a greater market orientation.

We concluded our fieldwork when we achieved theoretical model saturation, a point at which additional fieldwork appeared unlikely to change our model significantly (Glaser and Strauss 1967; Strauss and Corbin 1998). Finally, we prepared our findings and sent copies to all interviewees for member checks, both to help verify the accuracy of our representation and interpretation of each firm's experiences and to provoke any additional insights our informants might harbor (Belk, Sherry, and Wallendorf 1988; Lincoln and Guba 1985).

Results: A Four-Stage Model

Our analysis shows that the process of creating a market orientation occurs over four distinct stages: (1) initiation, (2) reconstitution, (3) institutionalization, and (4) maintenance. The stages are path dependent, and each stage includes multiple steps or activities, which we describe in this section.

Stage 1: Initiation

Given the vast literature suggesting the many advantages of a market orientation, investigations into the antecedents of a market orientation have overlooked what initiates market orientation change efforts. Our research finds that change is precipitated by powerful stakeholders recognizing a threat, who then create coalitions to plan and implement change efforts.

Recognition. The impetus for change was the recognition of an external threat. One dimension of such a threat was financial. All firms failed to meet financial performance targets. For example, Marshfield DoorSystems was losing millions of dollars on $53 million in revenue the year the company began its transformation. Some firms faced additional challenges. Competitors or changing technology threatened their business models, their industry leadership position, or their organizations' legacies. For example, having created the cellular phone industry, Motorola saw its leadership eclipsed by Nokia in 1988. Losing the leading sales position was a blow to the essence of how employees viewed themselves and their roles within Motorola as a world-class engineering powerhouse.

Within each participating firm, powerful organizational stakeholders (owners and/or senior executives) recognized these threats and began the process of transformation. The number of stakeholders initiating change varied across companies. In the cases of Marshfield DoorSystems and Alberto-Culver, a single executive initiated the process. At Harley-Davidson and Motorola PCS, a small group of powerful stakeholders began the process. In every case, stakeholders faced organizations that, for the most part, did not agree with their sense of how significant a change was needed or the nature of the change required.

Preparation. In the preparation step, dissatisfied and powerful stakeholders initiated backstage preparations for the transformation, spearheaded by a new leader. In some cases, the new leader was also the stakeholder who recognized the need for change; in other cases, the stakeholders appointed a new leader. Preparation activities included development of a guiding coalition based on a shared set of values and agreement on a broad plan for change. Coalition membership was based on two criteria: a genuine respect for the ability of every organization member to contribute to the organization's future and a belief that market needs should drive organizational change. For example, a Weyerhaeuser corporate vice president appointed Bill Blankenship president of Marshfield DoorSystems. Before officially taking over, Bill Blankenship recruited Jerry Mannigel, a Weyerhaeuser general manager who was well regarded by local management and the union and with whom Bill Blankenship had previously worked. Jerry Mannigel then recruited other coalition members, including a successful production manager who had previously left Marshfield DoorSystems in frustration.

Across firms, these guiding coalitions developed transformation plans that would engage the entire organization in the change effort. Plans focused on organization culture and process changes rather than on specific end-state goals or objectives, such as market share or return on assets. For example, four weeks after taking over at Marshfield DoorSystems, Bill Blankenship and Jerry Mannigel held a senior management meeting to outline a new set of values and norms expected of organization members. These included empathy, respect for others, collaboration, teamwork, and a market orientation. The plan for change included six process-focused change initiatives: (1) value proposition and value delivery systems, (2) communication systems, (3) safety, (4) continuous improvement, (5) organizational alignment, and (6) information technology.

In contrast, firms that were failing to create a greater market orientation did not have guiding coalitions based on shared values or agreement on a plan among people expected to create change. Instead, leaders relied on existing management structures, new hires, or consultants to create change, and their plans focused on improving their firms' marketing capabilities and cultures. Plans included creating new brand identities at BenefitsInc and EquipmentCo, hiring experienced marketers to infuse marketing into the organization at BenefitsInc, creating a corporate marketing council to advise and oversee business-unit activities at MediaCo, and marketing and cultural change training at MediaCo and EquipmentCo.

Stage 2: Reconstitution

Reconstitution is the organizationwide implementation of the plan for change. It comprises five steps: (1) demarcation, (2) value and norm development, (3) reconnection with the market, (4) removal of dissenters and hiring of believers, and (5) collaborative strategy.

Demarcation. A milestone in the change process occurs when the guiding coalition presents its plan to the entire organization. Within participating firms, informants uniformly mentioned the same milestone as the point at which people knew about the change program. Demarcation events were characterized by several common attributes: (1) They were public and open, and all organization members simultaneously witnessed the event; (2) the challenge to the firm was explained as authentic and compelling; (3) the plan for change and the values, norms, and behaviors the coalition sought to inure were clearly communicated; (4) the change process was framed as market focused, not marketing centric; (5) and the guiding coalition's actions were showcased in harmony with the desired culture.

The ceremony at Marshfield DoorSystems is an excellent example of the dramatic nature of such events. In January 1994, Bill Blankenship and Jerry Mannigel had been in charge for three months when they called a companywide meeting. People were worried. No executive team had ever called a companywide meeting. People entering the makeshift hall inside the plant noticed that something was more than amiss; it was odd, out of context, and blatantly ridiculous. Railroad crossing posts—lights and all—flanked the podium where Bill Blankenship was standing. An old-fashioned, steam-driven passenger train, replicated on a banner that was strung up above and behind the podium, ran 15–20 feet across and 3–5 feet high. Pasted on the locomotive was a picture of Bill Blankenship, as engineer, looking out the window. Clearly printed beneath the train were the words, "There's a new train in town." People's recollections of that day were consistent:

> I'll never forget the first meeting we had when he was introducing himself. He had this banner up,… [and then he said,] "There's a new train in town. You can get on the train or you can choose not to. If you choose not to, when the train leaves the station, you won't be on it."
>
> It's one of his classic lines. "There's a new train in town" Some people are going to come along for the ride. Some people are going to drive. Some people are going to decide to stay at the station. And some people—it will be decided—they will stay at the station. You did have options. It's your call. What do you want to do? Are you gonna come along? You gonna help us?
>
> The new train in town was really more about how we're all going to march to a similar step. Not necessarily the exact same one, [but] this is the direction we're going. We're going to do business differently. The analogy with the train is: You can get on the train. You can get run over by the train. Every once in a while, the train stops and lets some people off, and you have a choice: What kind of a passenger you want to be? Everyone agreed they didn't want to be run over by it. So most people jumped on board right away, but [the train] had to stop a couple of times and let people off.

The train analogy captured people's attention, and the companywide meeting marked the first time senior management publicly outlined the transformation plan. That everyone was informed simultaneously was noteworthy on its own. At the same time, management opened the organization's books to the union, so the union could review Marshfield DoorSystems's financial condition firsthand. More than eight years later, informants recalled the ceremony and related actions as a turning point for the company, when management unambiguously involved everyone to participate in the transformation of the firm.

EquipmentCo, MediaCo, and BenefitsInc did not have salient demarcation events. Instead, their firms communicated change efforts over time through formal channels; senior managers were notified first, followed by middle managers, and so forth. Informants recalled participating in specific workshops or training sessions, but they did not recall any overall plan for change, other than increasing organizational and/or marketing effectiveness. Finally, informants at the three firms that were failing to achieve a market orientation discussed how executives often acted in conflict with the espoused values and norms suggested by change efforts, and several senior managers publicly disagreed with official change efforts.

Value and norm development. The guiding coalitions perceived their existing firms' cultures as major impediments to organizational change and long-term success. Organization cultures before transformation were diverse, but they can be characterized as sharing the following attributes: a bureaucratic and internal focus; a reliance on historically successful approaches to solve new problems; highly structured routines parsed out by function and historical origin; employees identifying more with their function, job class, location, or other subgroup than with the overall organization; low levels of trust between groups within the organization; a lack of a common understanding of what the firm was trying to accomplish and how; passive-aggressive behaviors due to internal norms to be nice, coupled with the simultaneous use of covert methods to accomplish personal or group goals; and transactional leadership, defined by the extensive use of explicit rewards to manage employee behaviors.

Recognizing these challenges, the guiding coalitions attempted to create a new set of organizational values that, in combination with subsequent activities, would create a more market-oriented culture. The number and nature of these values varied across firms, from Alberto-Culver with ten values (and the mnemonic "HOT CC FIRST" to help employees remember them all) to Harley-Davidson, which communicated its values implicitly through a focus on customers and internal cooperation. Among participating firms, we identified six common cultural values. We list these in Table 1. Following Schein's (1985) organizational culture framework, we provide the basic assumption underlying each value and the behavioral norms emanating from each value as well.

As an organizational value, the market as the raison d'être provides common meaning for all organization mem-

TABLE 1
Values, Assumptions, and Norms of Market-Oriented Firms

Value	Assumption	Behavioral Norms
Market as the raison d'être	We come together as an organization to serve the market and make a living.	Every decision and action must consider how it affects the market.
Collaboration	Working together, we can achieve more, faster and better, than apart.	Work is done collaboratively by teams. Teams are jointly responsible for outcomes.
Respect/empathy/perspective taking	People are basically good and have reasons for their actions.	Consider the perspectives, needs, training, expertise, and experiences of others when reacting to or interpreting their actions.
Keep promises	To succeed, everyone must do his or her part.	Each employee is responsible for following through on commitments to others.
Openness	Honestly sharing information, assumptions, and motives allows others to understand and effectively collaborate with us.	Proactively and honestly share information, assumptions, and motives with others.
Trust	Everyone is committed to the same goal. Therefore, we can have positive expectations about their intentions and behaviors.	Trust that your fellow employees are telling the truth and will follow through on commitments.

bers and is broadly inclusive, creating one socially constructed in-group that encompasses all organization members. Such an inclusive in-group definition fosters positive attributions and behaviors among organization members and creates a cultural foundation for firms that allows for the ongoing adaptation of missions and visions in response to changing market conditions. In addition, we found the market as the raison d'être to be the central cultural value providing a rationale for the other five values of trust, openness, promise keeping, respect, and collaboration, all of which are required for an effective market orientation. As such, the combination of the six cultural values creates an organizational environment that is supportive of organizationwide collaboration, leveraging the experiences and capabilities of all members, creating a shared understanding of problems, allowing for the creation of more effective solutions to problems, and assisting in effective implementation of solutions through tighter collaboration.

The guiding coalitions used three methods to encourage adoption of these values and norms and, consequently, behaviors consistent with these values and norms. First, the guiding coalitions created change by explicitly modeling behaviors aligned with the desired values and norms. Second, they used rewards and recognition to encourage behaviors that were aligned with the desired values and norms, and they penalized (including removal from the firm) failure to adhere to the desired values and norms. Finally, they inculcated desired values and norms in organization members through participation in subsequent transformation activities.

Beginning with demarcation, the guiding-coalition members' behaviors and organization rituals played an important role in changing firms' cultures. For example, following the guiding coalition's leveraged buyout of Harley-Davidson from AMF, there were several symbols of management's intention to create a culture of trust and collaboration focused on the needs of Harley-Davidson customers. These included sharing detailed financial information with the union, as one informant explained:

> [U]ltimately, when you know how desperate the situation is and you truly need assistance—whether it's from the salaried workforce, or the union workforce, or your lending institutions, or your dealer network, or the government, or whomever it may be … to cut through the cynicism, I think you've got to show them the facts.… If you're really in trouble and you need someone to extend a helping hand, whomever that may be, you damn well better be honest with them. The way you're honest with them is opening up the books and sharing management responsibility, working with the union, [and] looking at the union as a fellow stakeholder.

Other culturally consistent symbols at Harley-Davidson included the elimination of assigned parking at headquarters; participation of senior executives in demo rides, rallies, and local rides (alongside other employees); one cafeteria at headquarters that everyone had access to; and the policy that *everyone* needed to obtain their personal motorbikes through the dealer network as any customer would.

The guiding coalitions also used rewards and recognition to create cultural change. The most prevalent recognition was interpersonal, with the guiding-coalition members showing approval, support, or disapproval of particular employee comments or actions in various venues, such as at Harley-Davidson demo rides, in Motorola PCS war-room meetings, and on Marshfield DoorSystem's factory floor. By demonstrating appropriate behaviors in front of the guiding-coalition members, employees signaled their adherence to change efforts, to which the guiding-coalition members reacted favorably by complimenting or paying attention to employees.

More public and permanent recognitions that transcended time and place were also used to engender cultural change. Alberto-Culver used public awards extensively, as one informant explained:

> Everybody goes into a room and the awards start getting handed out. The guys who are glued to their chairs and never getting up and getting an award with red faces are very, very identifiable. Then those that need a wheelbarrow to take all the awards back to their office are also very visible. So there is a separation between the haves and the have-nots. It is a very, very public thing.

Regardless of the mix of methods organizations used, we found that their effectiveness at creating cultural change varied directly with their "authenticity." Authenticity refers to the perceived congruency among management actions, rituals, symbols, and the values and beliefs espoused by the guiding coalition. When discussing the change process with the guiding-coalition members at Harley-Davidson, Marshfield DoorSystems, Motorola PCS, and Alberto-Culver, we found that they exhibited an almost evangelical passion and commitment to the change process and espoused cultural values. Their emotional commitment was unambiguous and entirely consistent with informant recollections.

For BenefitsInc, a firm that was not realizing a market orientation, we found no cultural change efforts. Conversely, MediaCo and EquipmentCo engaged in firmwide cultural change efforts, but these efforts were managed by an organizational development consulting firm (both companies used the same consulting firm). All employees were required to participate in cultural change training. However, informants perceived these formal cultural change efforts as insincere. For example, executives received three or more days of extensive training a year before the lowest-level employees received a half-day of training, contradicting the notion that all employees have something to contribute. Informants also noted that some executives openly questioned the value of the training, and other executives claimed to support the desired culture but did not follow the espoused norms. One MedicaCo informant summarized his frustration:

> I think there's this big disconnect. You have those things, the values and visions of the company from [the CEO's] mouth. But that's about as far as they fall. I mean, what are those things? I mean, what do they really mean?... [M]onth, after month, after month, after month of culture transformation exercises and then different [planning] initiatives,... that's where you create credibility gaps with senior leadership who are trying to lead. Because people say, "Wait a minute! *This doesn't sync up!* You tell me one thing but you're doing another."

Reconnection with the market. Organization researchers assert that cultural change requires changing organization members' underlying assumptions about how the firm works and what behaviors are appropriate (Schein 1985). Value differences between more effective and less effective organizations exist not in their labels but in what those values mean and the appropriate norms and behaviors tied to them (Argyris 1990). The same appears to be true for creating a market-oriented firm, the greatest challenge being to determine exactly which behaviors align with the market as the raison d'être.

For example, EquipmentCo invested heavily in cultural training compared with other firms in our study. Nevertheless, organization members were vexed by the ambiguity of what a customer focus meant and, thus, what behaviors exhibited a customer focus. A sales manager disclosed, "I'm frustrated because I don't think ... we *clearly understand* the market that we're trying to sell [our products] to.... We keep talking about this value customer. Who is that guy? What does he look like? You know? How does he act? I can't tell you.... My frustration really comes from I don't know what winning is. I don't know what winning in a value environment is. Because we haven't defined that. Tell me the 1500 guys that you want me to sell, and I'll sell 'em, but I don't know which 1500 right now."

Informants at Harley-Davidson, Marshfield DoorSystems, Alberto-Culver, and Motorola PCS recalled similar frustrations before their transformations. This suggests that an organizationally shared market understanding differentiates more market-oriented firms from those earlier in the change process and differentiates more market-oriented firms from the same firms before beginning the transformation process. Thus, to create a greater market orientation, firms create an organizationally shared market understanding, enabling organization members to determine appropriate behaviors that are consistent with the value of the market as the raison d'être.

Firms create shared understandings—or shared market schemas—by sending cross-functional teams into the market to meet with customers, channels, and influencers. Over time and across teams, those experiences coalesce around a shared understanding of the market. Teams explicitly represent this understanding in a market metaphor, such as "a Harley Rider" for Harley-Davidson. Teams then vicariously share their experiences with other organization members using stories and artifacts from the field, explaining the meaning encapsulated in the market metaphor. It is through this process that firms create a shared market understanding.

For example, as the first step in developing a value proposition at Marshfield DoorSystems, cross-functional teams were sent out to visit architects, contractors, and distributors. Informants described these visits as transformative. An informant with more than 20 years of industry experience shared the following:

> I went on a few trips.... We would ask contractors, "Why are things like that? Why is shipping on time important?" [It] was *really very interesting* for me.... Usually only sales guys got to see the customers.... I got to understand some of the issues, like delivering in the right week is what it's all about. Like issues with doing remodeling and having to have service elevators scheduled. [The contractors] need precision in big cities like Boston. These are considerations I never knew about before. It's important

> because now we're talking about the end-use customer. That end-use customer wants it in there at a certain point in time. This has a great impact on the architect [and the] contractor that specified your door and the end-use customer who is going to pay for it. If you do it well, they'll be back. If you don't do it well, they will remember that,… and of course, it all has to be correct.

The informant was aware that delivering products to specification and on time was important, but meeting distributors, building owners, general contractors, and architects created a new sense of urgency and importance around these expectations. The informant developed an appreciation for *why* these things were important and the impact of failing to meet such expectations. When he discussed his firm's value proposition, he relayed these personal stories, providing context and understanding for the value proposition. This allowed a listener to appreciate the perspectives of the people he met, leading to empathy for their situation.

During the 1980s, Harley-Davidson initially sent employees into the field to staff events because they did not have the money to hire outside people. Soon after, Harley-Davidson began to realize benefits similar to those at Marshfield DoorSystems, and eventually, all employees were *expected* to work in the field to understand customers and their needs (e.g., Fournier et al. 2000). Motorola and Alberto-Culver sent smaller groups of people into the market, who similarly brought back market insights to create organizationally shared understandings in their firms.

At Motorola, the Global Design Planning team was responsible for managing fieldwork as part of the product development teams' uncovering consumer needs and trends. One team member explained, "The [Global Design Planning] team does behavior research…. Behavioral research uses a lot of social science methodology—anthropology methodology." Using a project example, the informant explained, "No one knew what [a leadership phone] meant, other than the definition we had with StarTac…. [So] we actually went out and did interviews and made a research brief. We went out and interviewed the leaders, the cutting-edge people."

To share this field information, team members brought back stories and artifacts, including video recordings and pictures. This allowed others to experience the market vicariously and build a shared understanding, much like the employees at Marshfield DoorSystems who heard stories from returning colleagues. A guiding-coalition informant, focused on design, summarized the transformative nature of this process at Motorola:

> This is a journey of enlightenment, right? What is enlightenment? Enlightenment is when you truly understand the true cultural points of resonating with what their needs are. What does that mean? That means there's a huge installed base of knowledge within the culture about the world it shares. There's a constant feeling of that knowledge with ethnography that many other points [miss]. It's that point of enlightenment. It's also at the point where you're truly innovating cross-functionally, not only with technology and design but also with the business that you're creating…. Enlightenment, in that context, is a higher art of brand building and value creation.
>
> It comes back down to this notion of collaboration and shared understanding as being really fundamental; you have to have opening the kimono itself be a core function of the company and build a shared vision of where you're going, share the enthusiasm, and that requires a level of intensity and communication that's really never to be underestimated.

In addition to product-specific research, Motorola PCS created a brand compass metaphor that highlighted four attributes across market segments: communication, business use, social use, and price. The brand compass was a triangle, with communication as the base, business and social as the sides, and price increasing from entry level at the base to premium at the top of the triangle. Pictures of five archetypal consumers were superimposed on the triangle, symbolizing five broad consumer segments: everyday communication, easy business, corporate business, networked entertainment, and personal style. Although we did not observe creation of the brand compass, its use in almost every conversation within Motorola PCS was striking, particularly the stories people offered to contextualize each segment. All products were first defined within the brand compass metaphor and then elaborated on the basis of their specific attributes, such as corporate e-mail access or fashion/style consciousness. The brand compass allowed people to discuss various handset and service initiatives within a common understanding of what they were trying to accomplish collectively.

Reconnecting with the market is not defined by fieldwork alone, but also by how such experiences are shared in the field from multiple perspectives and then shared within the organization to create shared meaning. Informants from each of the firms that were failing to realize a market orientation recalled sporadic field visits with customers conducted by functionally homogeneous teams that did not share the results internally. A BenefitsInc informant recalled,

> That reminds me, the first boss I had here believed that people should leave the office. He always thought the people at [divisional headquarters] really never left the office. So he wanted people to go out and do customer interviews. [We found an academic specializing in customer visits and had] an internal interviewing training program to teach people who are not market researchers how to conduct an interview…. [My boss] got all the people in marketing … and some sales people too,… and they went and interviewed customers…. There were teams, three or four at a time…. It worked out pretty well,… but the follow-up through some type of methodical "what did we find" didn't really ever get done, at least not that I know of.

Removal of dissenters, hiring of believers. Although initial cultural transformation efforts are sufficiently effective early in the transformation process, at some later point, coaching cultural dissenters to change is abandoned and removal of such dissenters becomes the focus. Simultaneously, organizations modify their recruitment processes, henceforth requiring that individually held values of new members closely match those of the organization. These changes suggest a higher-level change from creating a new culture to preserving and nurturing an existing culture.

Experiences at Alberto-Culver are revealing. At one point, 75% of President Carol Bernick's direct reports were new. Some were part of her initial guiding coalition, subsequently elevated to new positions, and others were new to the organization, recruited on the basis of functional competencies and ability to fit into the culture. As one informant recalled,

> [Bernick] made some changes fairly early on in the marketing area. She made a change fairly early on with her financial guy.... So those two areas changed sooner than later. I can't remember exactly on the sales side. I think she changed it once and then changed it again with Dick Hynes. So I mean, she built her team. The most recent change has been in [research and development] ... a year and a half ago. She really feels strongly, and I absolutely agree with her that she needs people who really understand the culture and who operate within the culture.... [O]bviously you've got to perform in your function, but you have to understand and support and drive the culture into the organization.

In conversations with the newer executives, it was clear that their values were aligned with the firm's values. The Alberto-Culver culture played to new executives' strengths and fit what they were looking for later in their careers, following successful careers with much larger, bureaucratic organizations. When we asked newer executives about collaborative team meetings with invitees from across the company, they were often puzzled by the question and said they had always worked that way. Thus, apart from removing people who did not fit the new culture, much of Alberto-Culver's success at becoming more market oriented was attributed to bringing in people who already believed in the culture and were able to add to it through their unique experiences and competencies.

Regarding firms that were failing to realize a market orientation, there was no comparable removal or hiring of people based on cultural fit. On the one hand, because their firm cultures appeared much weaker, it is difficult to imagine how they could make personnel decisions based on cultural fit. On the other hand, it appears that the lack of consequences for defying the desired cultural norms impeded cultural change and thus resulted in weak overall organizational cultures.

Collaborative strategy. At this transformation point, organizations collaboratively address tasks in a more market-oriented manner, precisely because they agree on what the market is, what its unmet needs are, and how to work together to meet those needs. Furthermore, the collaborative development of strategy leverages a firm's collective knowledge and capabilities, leading to more creative and realistic strategies (i.e., they can be implemented). In addition, the organizationwide comprehension of strategic goals and their relation to the market provides context for why specific tasks are important, increasing the likelihood of success. This collaborative process also leads to higher levels of commitment to the strategies and the firm, which again increases the likelihood of success.

Indicative of this process, a Motorola executive explained the change in developing strategy as paradoxical; there were more people involved than ever before, but it was more productive and definitive than pretransformation efforts:

> What you end up doing is creating a camaraderie that ... may not have any effect at all on the definition of the product, but ... you've got a team that can then work through normal process, which is more condensed. It's more focused.... What's fun is ... you make sure that you're open-minded to those big sessions—that you could [reveal and learn new] stuff,... and that's a really wonderful surprise and great things can happen.... At one point,... we had 300 people involved in developing the strategy, and it was kind of like, "What the hell are you doing?"... but it built a portfolio that was focused and clear and driven when we got done with it.

Similarly, Marshfield DoorSystems informants explained that their new culture and value proposition made strategy development and implementation much easier because, in the words of one informant, "There was *one* picture that we all focused on, and everyone *knew* it!" A shop floor employee recalled, "Basically, when Bill came, he turned the mill over to the *people* and said, 'This is your baby—make it work!'... [Today,] the union and the company are working *together* instead of *against* each other."

Although the process of developing and implementing market-based strategies was more efficient and effective across firms because of a new shared understanding of the market, this shared understanding did not replace traditional marketing research. As a Harley-Davidson marketer observed, "Market research just meant more.... [I]t made it much easier when presenting findings or ideas within Harley-Davidson. People would say, 'Oh yeah! I remember this one guy.'"

Conversely, this new collaborative approach changed the role of formal organization structures. From this stage forward, collaborative cross-functional teams conducted the vast majority of work; any individual could be on multiple teams, and no two teams comprised the same members. This concept was exemplified by a Motorola informant's observation that "one of the other things Mike says 50 times a week is, 'Make the matrix work. I know we're a matrix organization. I did it on purpose.... I know it's not easy.... Get your ass down there and figure it out.'"

Without the benefit of a shared market understanding and agreed-on norms for behavior, the three firms that were failing to achieve a market orientation were demonstrably frustrated in their marketing strategy efforts. Introducing the new brand identity at BenefitsInc was laborious as a result of different conceptions of what the business was versus the proposed brand promise. After the new brand identity was implemented, senior managers believed that it would lead to organizational change, but a member of that effort confided that there was now "this sense of imbalance in terms of what the marching orders are versus what can be realistically achieved. I think, underneath it all, employees understand that big difference, and if they understood that, 'Hey, you know what, there's a realistic mission here, and we're quick to accomplish it,' they'll do it. Otherwise,...

you're told, 'OK, it's D-day, scale that cliff.' [Then the employees say,] 'Well, with what? What kind of backing do I have?'"

Even with the cultural transformation workshops and assistance from outside business consultants, MediaCo and EquipmentCo informants shared a frustration with strategy development, and no one seemed to understand how it all fit together. A MediaCo informant commented, "There's a [strategic] plan, and it's completely divorced from what we're doing.... [A]lthough we suffer from that like everyone else, I've never seen anything ... so completely decoupled [from reality].... We don't have a planning process that actually links our [strategic] plan to our operational plan. In fact, we don't have operation plans; we have budgets." Similarly, an EquipmentCo informant shared, "[W]e were talking yesterday about the planning process in general. We got in these kinds of debates of when we have to have our plans due, and I was like, my team, it's very hard for us to develop our plans before the [product] centers have their plans in place, because what we need from them is, 'What target markets are you going after? What segments?... What kind of market share growth do you want...? What are the key value drivers...?'"

Summary. The reconstitution stage consists of taking the new concept for the firm public with a demarcation ceremony. After the plan becomes public, the effort turns to value and norm development, reconnecting with the market, personnel changes, and the development of a collaborative strategy. As a result of these efforts, the reconstitution stage creates the culture, understanding, and processes that lead to the organizationwide generation of, dissemination of, and responsiveness to market intelligence described. Elements of this process or ethos have been discussed by other researchers (e.g., Allen, McQuarrie, and Barr 1998; Gouillart and Sturdivant 1994; Kohli and Jaworski 1990; Shapiro 1988). For example, others have advocated cross-functional customer visits (Allen, McQuarrie, and Barr 1998). Our analysis shows how these elements are woven into a process. Moreover, we show that successfully creating an organization that behaves in such a way depends fundamentally on new values and norms that are reflected in organizational learning, personnel decisions, and the collaborative strategy development process.

Stage 3: Institutionalization

By the time firms reach Stage 3, they have undergone dramatic, fundamental changes. These changes, however important, remain informal to a significant degree. That is, the organizational culture and resulting behaviors have changed much more dramatically than the formal organization structure and supporting attributes. In Stage 3, organizations institutionalize these changes. The speed and effectiveness at which organizations successfully institutionalize these changes depends heavily on their experience in prior stages. Firms that realize greater levels of market success develop increasing comfort working in a market-oriented manner, accelerating their institutionalization. Simultaneously, the guiding coalitions create explicit symbols, artifacts, and rituals to retain and further encourage a market-oriented culture, characterized by (1) formalization, (2) alignment of rewards, (3) indoctrination and training, and (4) power shift. At the time of our fieldwork, Marshfield DoorSystems, Alberto-Culver, and Harley-Davidson had progressed through the institutionalization stage.

Formalization. The guiding coalitions explicitly formalized their organizations' evolving market-oriented cultures through symbols, rituals, and artifacts. These included organization structure changes at Harley-Davidson, the implementation of high-performance work systems at Marshfield DoorSystems, and the explicit defining of cultural values at Alberto-Culver and Harley-Davidson.

In 1993, Harley-Davidson moved to a circular organization structure, formally institutionalizing its collaborative culture. The functional leadership group was composed of Harley-Davidson's top management with three subgroups: create demand circle, produce product circle, and provide support circle. Even with a history of participation and collaboration, a Harley-Davidson executive explained how the circle concept further raised expectations and opportunities for collaboration over time:

> Circle meetings are pretty functional, although I don't think they always have been. I think we had to find that fine line about what is *circle-worthy* and what is not. There was a time ... I was coordinating some function and the circle ... sat there and argued about how long the cocktail hour should be. OK, so you've got a bunch of [vice presidents] sitting there arguing about how long the cocktail hour should be. So you could fall into a trap. I think they had to find their way about how detailed do we get in our understanding.... In that respect, it's important to recognize what do you need other people's support and input on and what don't you. What is within your own function? There is no book that you could have followed that said, "This is a circle thing, and this is not." It's a judgment call. When everybody came in, it was just that—some people wanted to know everything about what everybody else is doing—so for everything, "Well, the circle should decide that," and other people said, "Um, I think that's my job to decide that...." When Rich [Teerlink] put us all in the circles, along with Jim Paterson,... they just said, "Here you go. And here's the philosophy around it." So I think there's a lot of struggling to figure out where that is. That's where a lot of trust and communication has to happen. You've got to assume that everyone is doing their job and doing it well. Otherwise you could make everything a circle issue.... So I think we still struggle with that sometimes, but we'll catch ourselves now.... So we had to find the point where there's constructive interaction and where there's just simply meddling. Meddling is not productive. But it is impossible to live in a silo because we're so integrated with each other There's very little that I do that somebody else doesn't feel, and vice versa.

Alignment of rewards. The intrinsic motivation from serving the market and working collaboratively toward common goals appears to be the primary driver for employee efforts in earlier stages. Greater levels of intrinsic motivation are created as organizational cultures become stronger and firms begin to experience success. However, at a point when employees feel empowered and the organization is successful, employees perceive working harder as primarily benefiting the organization's owners and top management. This creates perceptions of injustice due to the incongruity between the culture and the allocation of rewards. Within participating firms, the implementation of group-based variable compensation schemes addressed these concerns and further institutionalized a market-oriented culture.

For example, despite increasing levels of success in the late 1980s and early 1990s, Harley-Davidson's management perceived a growing tension between the firm's culture and its formal incentive structures, which had remained largely unchanged since the early 1980s. Teerlink and Ozley (2000, p. 145) note, "Employees argued that, at the end of the day, the 'real' work of Harley-Davidson was the work that the company recognized in a paycheck. This 'touchy-feely' stuff seemed okay,... but the organization would resist the kinds of behaviors that were being proposed until the company recognized and rewarded those behaviors."

To address these issues, the company first implemented a new personnel development and evaluation system, matching individual expectations and performance directly with the company's mission, values, and business processes. An executive commented, "This was a vision of the way people needed to be engaged in an organization ... that allows for theoretical alignment of an individual's job with the long-term direction of the company" (Teerlink and Ozley 2000, p. 106). This alignment also provided flexibility in adapting to changing market and organizational conditions. Employees were expected always to do what was best for the organization under the broad umbrella of its values and objectives, without having to make trade-offs between individual compensation and the organization's welfare.

In parallel, everyone's compensation was altered to include a variable component that corresponded with Harley-Davidson's success (or lack thereof). Management defined performance metrics at the business unit or plant level (the largest collaborative group with control) and progressively expanded the program to include all salaried employees by 1990 and all employees by 1993. Variable compensation bonuses were paid to everyone or no one on the basis of the goal levels achieved.

Indoctrination and training. Over time, the saliency of shared understandings from the reconstitution stage degrades. More significantly, as firms recruit believers and remove dissenters, an ever-growing number of new employees lack the shared experiences to understand assumptions underlying the culture. In addition, employees and managers often lack the knowledge or skill levels required to meet escalating expectations as firms become even more market oriented. To address these challenges, firms institute indoctrination programs for new employees and ongoing training programs that reinforce and refine the culture through skill development for current employees.

For example, in 1990, Harley-Davidson instituted employee orientation programs. An informant recalled, "When I came here there was an awful lot of formal training.... Within a few months of me being here, we had a new employee orientation. It was like the first one they had had. It lasted a half a day. And it was everybody who had started over the last six months. Now they do it, it starts every week, and it lasts three days, but it was a smaller company then." In addition, the informant received a copy of *Well Made in America: Lessons from Harley-Davidson on Being the Best* (Reid 1990), which documents Harley-Davidson's transformation from the early 1980s: "I read that one too before I started. They gave it to you before you started. Or at least they gave it to me."

Around the same time, the firm launched the Harley-Davidson Leadership Institute to reinforce the culture, provide a broader understanding of the firm, and familiarize employees with tools for working within the firm. Employees spoke enthusiastically about the Leadership Institute. One informant grabbed a Leadership Institute binder above his desk and, while thumbing through it, explained, "Have you seen this? Here are our values. Here's our issues, our stakeholders, our visions. You've heard these before, right?... So I won't recite them to you—because I can. And then you've got mission and operating philosophy, our objectives, and then my area's functional strategy.... [W]e'd spend time on what is this and what is [that] and how does it impact us in our area and what are you going to do about it? Make sure we're all connected here." Another informant commented, "[Such training] is probably more valuable today, to allow people to interact and understand events and come to some philosophy that was practical for them. It takes that philosophy and makes it practical for them. 'What does 'tell the truth' mean to you? What does the vision statement [mean to you]? What are stakeholders? A shareholder is a stakeholder—no more, no less than some of these other stakeholders.' To really get those issues out on the table and have the discussions about them."

Power shift. As firms realize higher levels of market orientation, the power to make decisions and act on behalf of the firm shifts from the guiding-coalition members to all organization members. Employees are responsible for making decisions and acting appropriately within the agreed-on framework of each firm, as well as ensuring everyone else is doing the same. Within participating firms, the guiding-coalition members viewed these behaviors as evidence of a successful transformation.

Indicative of this change, a Marshfield DoorSystems informant commented,

> We had [a reengineering blitz] that took place about ten days ago....[T]here was a group in the shipping department. They were doing a blitz on the palletizing and packaging area. We had seen a large number of claims, complaints, and frustrations about products not being packaged properly. Our biggest problem was distributors complaining that "It wasn't packaged how I ordered it." So we had this blitz going on,... hourly people along with the supervisor.... As part of this blitz, they sat down and interviewed one of our salespeople and asked, "Why is this such a problem?" Something *magical* came out if it! The salesperson said, "You've got to understand, the distributors are paying for this special handling, and they're not getting what they're paying for." Then one of the hourly people said, "Wait, let me see. Am I to understand this right? We know they're paying for palletizing, but you mean some kinds of packaging cost more than others?"

> The salesperson said, "We have twenty different ways of palletizing, and there's twenty different prices. Whatever you ask for has different prices." The people in palletizing didn't know that! They said, "Really? We just thought, some just wanted it this way, [and] some [others] wanted it that way. A lot of times we're busy, [so we figure] what the hell is the difference? I'm doing the next pallet; I'll do it just the same...." [So the team realized,] "Wait a second! We're creating value! They're willing to pay for this!" And you keep these learning things going on. So we had two different schemes: the sales people wanted to try more options; shipping wanted to get it down to [one option]. They didn't realize the two were causing problems for each other. The salespeople didn't understand some of the associated costs, even though they were charging differently; it wasn't appropriate relative to what the costs really were. So the solutions came out and the rest of that week they spent their time understanding every single one of the options. This is the pure cost of it. So we have a mark-up or discount and everybody knows it. Now they know why, so when they get a request for palletizing in a certain manner, [they understand that] this is what the customer wants.

In addition to decisions about how to serve the market better, employees are expected to enforce the culture by pointing out other organization members' cultural infractions. In other words, the power used by the guiding coalition to guide the transformation is distributed progressively among all organizational members to maintain that culture. For example, a Harley-Davidson informant explained, "We know when we have permission to do things and call people on things.... The issue is that training here gave us permission to call people on things; it gave us the community which we work in;... it's not a personal attack, but part of getting things done." Similarly, at Marshfield DoorSystems, we observed a union member challenge executive behavior regarding the cultural value of openness during an all-company meeting, commenting, "We used to get communication all the time. We used to have meetings every morning.... I think there is an issue of communication here....What are the monthly profit numbers? Shipping on time? We used to get this weekly." Not only was this acceptable employee behavior, a senior executive apologized for the oversight, promised to fix the problem, and suggested that the employee follow up if no action was evident.

Summary. The institutionalization of a market orientation requires formalization of organizational structures and processes, alignment of rewards, and cultural indoctrination through training. These changes institutionalize a power shift from the guiding coalition to the larger organization. With the exception of the power shift, previous research has focused on similar formal changes as a way to develop a market orientation (e.g., Day 1999). We find that, indeed, these changes are central. However, unlike previous studies, we find that these actions follow a more fundamental cultural shift. Thus, formal changes institutionalize and reinforce earlier cultural changes. Furthermore, the natural consequence of a strong culture is that it relies on individuals distributed throughout the organization to ensure cultural enforcement rather than relying on power concentrated within an elite guiding coalition, as was the case for instigating cultural change.

Stage 4: Maintenance

A curious phenomenon occurs after organizations achieve a greater market focus: Employees develop increasingly divergent interpretations about why their firms changed and the attributes responsible for their firms' success. Employees who join firms after transformation voice the most significant interpretation variations. Organization members who participated in the transformation also offer variations, and the degree of variation is positively correlated with the amount of time since the transformation.

Specific individuals or firm attributes are the most common explanation for a firm's market orientation and success. Individual-focused interpretations create a mythological figure or group of figures out of the guiding coalition's leader or members. Such interpretations transform the messy, progressive, and confusing process of collaboratively becoming market oriented, engaged in by hundreds or thousands of people, into the more appealing root metaphor of a savior-driven transformation (e.g., Turner 1975). Similarly, attribute-focused interpretations reconstruct history to fit current circumstances and enhance the inherent capabilities or importance of one of three causes: mythological figures, collective organization members' ability to face adversity, or pure luck. Such attribute-focused interpretations offer much simpler explanations than the actual series of changes responsible for firms' transformations. Although more easily comprehended, such attribute-focused interpretations create non-market-focused suggestions for how to handle future challenges, such as recruiting or revealing a new mythological leader, bearing down and working harder, or hoping for good luck.

Given these alternative explanations for market orientation and performance, organizational leaders develop processes and model behaviors in an effort to maintain a market orientation. Three processes reinforce a market-oriented culture: cultural screening of new members, culture maintenance rituals, and ongoing market connection activities to update market schemas and validate market-oriented process schemas. Firms also exhibit two characteristics protecting their market-oriented cultures from outside influences: Cultural "flame keepers" approve organizational changes based on cultural consistency, and there is a vigilance against management fads and fashions.

Cultural screening of new members. Alberto-Culver relied heavily on personal interactions to identify new employees, with Carol Bernick directing executive recruiting and approving managerial hires. Similarly, Marshfield DoorSystems relied on employee referrals to identify recruits who fit the culture, and senior executives were involved in every hiring decision. With a workforce of 8000 employees and incredible growth, Harley-Davidson screened new members using a prescribed process. To maintain a culture of "Harleyness," Harley-Davidson's recruiting process was extremely rigorous, focusing primarily on a cultural fit. In the words of one informant, "[W]hen you bring people into the organization that don't understand the tradition, your product suffers terribly. The neat thing about Harley-Davidson is we've managed to keep people coming in who understand the tradition. So the culture at the Motor Company absorbs them because they're ready to be absorbed and there's no conflict."

Culture maintenance rituals. Firms couple screening of new members with cultural indoctrination rituals that provide a historical context for how firms progressed to their current states and cultures. In contrast to previous indoctrination efforts, seasoned employees also partake in similar rituals, reminding them of the history and assumptions underlying their market-oriented culture.

For example, an executive informant explained that Harley-Davidson was considering some type of cultural tune-up program for all employees:

> [S]ome people have gone through it, and you have to be aware of that and do it as more of a refresher rather than relive the whole experience. So how do we do that? I think that's what we, as the senior leadership, are going to start talking about. I know that [another executive] has a focus on that—that we really need to start looking at that business process model and make it a refresher so it's not just the words on the wall.

Ongoing market connections. As organizations and markets evolve, cross-functional field visits appear to be crucial for maintaining a market orientation. Although all three firms conducted surveys and focus groups and performed other research activities to monitor the market and elaborate on their shared schemas, they also performed ongoing or periodic cross-functional field visits to ensure that their organizationally shared market understandings were up-to-date. As an Alberto-Culver marketing executive noted, "We do workshops around it—and I'll sit through some too—of the stuff in the field. I mean, I don't do the interview or stuff, but I'll certainly [be there]. Have to see some of the stuff with my own eyes. As much experience as I have in this category, it continues to change and people's beliefs change. So you have to stay with it and understand it, because sometimes my experience can be detrimental. I have a lot more baggage than most. I have to understand when something's outdated and when things change. So I try to get out there as often as I can."

Marshfield DoorSystems was updating its composite division's value proposition during our fieldwork. As with the original value proposition work, the firm was sending cross-functional teams into the field to interview buyers and so forth to develop a new value proposition based on their shared experiences.

With demo rides and rallies, Harley-Davidson had the greatest opportunity for and level of ongoing market connections. A company veteran commented,

> Even when I go to rallies now—I hadn't been to a rally in eight years when I signed up for Daytona this year when I met you—I came back with a greater sense of contribution to the business, visioning, and focus than I had before, and as [a high-level manager/executive], I'm hearing our needs, but it's good to get those needs either validated or repositioned hearing it from the customer.... [Regarding dealers], I've had my project managers going out in the field to understand what could/will collaborative inventory management [deliver]. They see customers. They see stocking in the dealerships. Our people see more about our business because we drive it by processes.... It's not just checking a box; it's driven by the market. I don't have to go to demo rides anymore, but I just wanted to.... Every opportunity you get to leave these four walls, you always learn something. I also go to a dealer almost every other month. It impacts how I direct resources, how I get projects approved, how I run projects, everything.

In addition, the company's program encouraging employee motorcycle ownership provided a venue for employees to experience dealerships and how Harley-Davidson's support programs affected end users. One informant explained,

> We encourage employees to own and use the motorcycles. Another point that I think is interesting about Harley is that we don't have an executive motorcycle fleet, but we do have a program that encourages people to buy a motorcycle. But the cornerstone of the program is that you go through the same thing a customer does. You have to go to a dealer; you have to get in line with everybody else. You've got to order a bike. I mean everything gets handled;... you make the best deal you can get. So the employees, in order to use the product, have to have the customer experience. We're not insulated behind a privileged fleet of bikes.

Cultural flame keepers. As firms adapt to changing market needs, employees are continually faced with procedural and other choices that could affect firms' cultures. Addressing the cultural impact of such changes rests with a small number of cultural flame keepers within each organization, whose implicit or explicit role is to ensure that procedures or programs adopted in an effort to improve the organization do not negatively affect the organization's core market-oriented culture.

For example, during a Marshfield DoorSystems senior staff meeting, the group discussed designing a system to track and reward employee suggestions for process improvements. A few executives had recently visited another company with such a system and were adamant that Marshfield DoorSystems should adopt the identical system and require all workers to participate. One of the original guiding-coalition members questioned whether such a system was consistent with the firm's core values. Following an elaboration on the cultural differences between the firms, the executive summarized his concern, "[Company X's] firm culture won't work here at Marshfield where there is a culture of trust, where you treat people like adults." All participants acknowledged the point and resolved to develop a more culturally appropriate system or not to implement one at all.

Vigilance against management fads and fashions. Another characteristic of firms maintaining a market orientation is an organizationwide skepticism toward management fads and fashions. Similar to the constant referencing back to shared market schemas regarding the appropriateness of actions, employees reference their shared market and process schemas before considering changes to any process suggested by management fads and fashions.

For example, during a discussion about customer relationship management (CRM) systems, a Harley-Davidson informant talked about the abundance of disparate customer databases available from dealers, Harley Owners Group, demo rides, registrations, and so forth. However, tying them together was not necessarily the right thing to do for Harley-Davidson, as the informant explained:

> Currently, each dealer has a great database of customer information. The dealer has it, not us. We have great databases too that they don't. The goal of the future would be to find a way to share what we know about our mutual customers and enhance the dealer's ability to sell and focus on more direct customer relationships, as well as us coming up with the right marketing programs…. The point is that we have a lot of data, but what do we do to turn it into actionable information,… and that goes back to a larger strategy of "What is our CRM strategy? How do we want to use customer information?" That's the biggest thing, and that's yet to be determined. That's the age-old question; one of the things that I've written down is I hear people say: "The real power of Harley-Davidson is the power to market to consumers who love the product." I mean, who else can say that? We're singing to the choir, right? People love us. What a job! We get to market to people who love us! I mean, car manufacturers can market to me. I don't love them.… The tough part is, if it's true, if the power of the Harley-Davidson brand is due to the power that we get because we market to consumers who love us, the question is, "Should we change that? Why would we get closer? Why would we appear intrusive? Why would we be farther away? Why would we change what we're doing?" So at other companies, they're now struggling with what is their [CRM] strategy; you know, you read it in every document you pick up: CRM, CRM, CRM.… Our people love us already. So the challenge here … is they love us already, we don't want to be intrusive.

Discussion

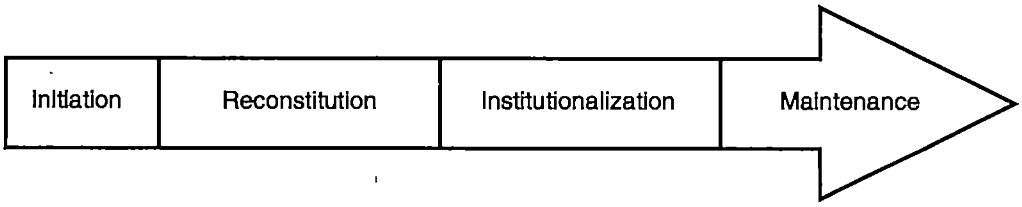
Our research shows that the creation of a market-oriented firm involves several interdependent changes at the individual, group, and organization levels that occur over several years. The four stages of that process appear in Figure 2, which shows that each of the stages has a distinct character and is composed of several steps. Stage 1 focuses on planning for change by a small group of powerful executives. Stage 2 is characterized by guiding-coalition efforts to build a consensus for change, based on new organizational values and norms and shared market experiences. If successful, the firm moves to Stage 3, in which it institutionalizes changes in organizational structure, processes, and formal rewards. In Stage 4, having achieved a greater market orientation, the organization creates mechanisms to sustain it.

Creating a Market Orientation Through Cultural Change

Our analysis, as reflected in this model, illustrates that creating a market-oriented organization is essentially a process of cultural transformation. Guiding coalitions create cultural change by inculcating organization members with market-oriented values and providing them with transformative market experiences. Cultural values are initially inculcated through management actions that exemplify desired values and through the recognition and rewarding of behaviors that are aligned with desired values. These management actions and rewards bear much closer resemblance to culture change levers that focus on intrinsic motivation, group dynamics, and social acceptance (e.g., O'Reilly and Chatman 1996; Schein 1985) than the teleological mechanisms of management direction and extrinsic pay for performance, as existing market orientation research suggests (e.g., Jaworski and Kohli 1993).

Our analysis reveals that creating cultural change requires direct or vicarious market experience, providing organization members with shared meaning and purpose. We find that organizationwide sharing and making collective sense of market experience is responsible for effecting a change to a market-oriented culture. The cross-functional, organizationwide nature of such activities creates an organi-

FIGURE 2
Process of Creating a Market Orientation



Initiation	Reconstitution	Institutionalization	Maintenance
1. Recognition 2. Preparation	1. Demarcation 2. Value and norm development 3. Reconnection with the market 4. Removal of dissenters, hiring of believers 5. Collaborative strategy	1. Formalization 2. Alignment of rewards 3. Indoctrination and training 4. Power shift	1. Cultural screening 2. Culture maintenance rituals 3. Ongoing market connections 4. Cultural flame keepers 5. Vigilance against fads and fashions

zation affiliation around serving the market as the most salient and attractive association for all employees. In other words, shared market experiences provide shared understanding and meaning to the cultural value that differentiates a market-oriented culture, namely, the market as the raison d'être. Although previous research has investigated the relationship between shared meaning and performance (Hult, Ketchen, and Slater 2004; Moorman 1995), our findings support the notion that shared market experiences contribute to a larger cultural change process to a market orientation (Narver, Slater, and Tietje 1998). Our findings also contrast with prevailing cultural change models in the organizational development literature (e.g., French and Bell 1999). Specifically, we find that firms that rely on organizational development consultants to facilitate change through behavioral interventions but do not reconnect with the market to create shared meaning and purpose do not realize a greater market orientation.

Beyond shared meaning and purpose, market experience provides opportunities for functionally or otherwise isolated employees to engage in perspective taking with others. As a result, employees engage in more generous attributions toward colleagues, facilitating increases in the cultural values of trust, openness, respect, promise keeping, and collaboration. These consequences parallel findings on perspective taking (e.g., Galinsky and Moskowitz 2000) and have been suggested as market orientation antecedents (interdepartmental connectedness and lack of conflict) in cross-sectional research (Jaworski and Kohli 1993; Kirca, Jayachandran, and Bearden 2005). Thus, within a larger process of creating a market orientation, our research offers insight into how firms create cooperation and reduce interdepartmental conflict previously not addressed.

Overall, however, our model for creating a market orientation differs considerably from the existing literature, which proposes three antecedents to a market focus: top management focus, interdepartmental cooperation, and reward systems (Jaworski and Kohli 1993; Kirca, Jayachandran, and Bearden 2005). One implication of this view is that actions taken to increase the level of any one antecedent will lead to a greater market orientation, all else being equal. In contrast, our results reveal a much richer, more complex process of organizational change. Our analysis shows that, indeed, top management focus, interdepartmental cooperation, and reward systems are important factors influencing organizational change, but the sequence of these factors in a larger process is essential to successful organizational change. According to our analysis, top management focus is required to begin and guide the process. Empowered stakeholders recognize a threat and begin the process of forming a coalition to transform the organization. Cooperation among functions is essential, but it requires top management support, common experiences, and a shared meaning structure to be effective. Organizational rewards are important as well, but our analysis reveals that this is the case only after top management support has been created, the challenge to the organization has been presented, the organization has been reconstituted, and an organizational consensus has been created for formalizing a new, successful firm with largely new values and purpose. Thus, rather than being three independent antecedents of change, top management focus, interdepartmental cooperation, and reward systems operate in a particular sequence as part of a larger, more complex cultural change process.

Market Orientation

Our findings offer new insights into the nature of market orientation, the role of organizational learning, power within the organization, and what sustains a market orientation. Most fundamentally, our results offer a new perspective on the debate about whether market orientation is simply a set of behaviors or a culture. Researchers advancing a behavioral perspective suggest that the activities of a market orientation are separate from organizational culture (e.g., Deshpandé and Farley 1998; Kohli and Jaworski 1990), whereas others suggest that it is an organizational culture that encourages such behaviors (e.g., Homburg and Pflesser 2000; Moorman 1995; Narver and Slater 1998). Our analysis provides further support for the cultural perspective. For example, as does Moorman (1995), we find that organizational culture plays an important role in organizational information processes. More important, our analysis provides new insight into the nature of the cultural values on which a market orientation is based. We find that firms creating a market orientation embrace six cultural values: trust, openness, keeping promises, respect, collaboration, and viewing the market as the raison d'être. The first five values encourage individuals to act as a cohesive whole in addressing the market. The market as the raison d'être provides individuals a rationale to collaborate toward a common purpose, thus supporting and reinforcing the other five values. These values are developed early in the transformation process, and they provide a foundation for changes that follow.

These cultural values are the basis for market-oriented behaviors, namely, the generation of, dissemination of, and responsiveness to market intelligence (Kohli and Jaworski 1990). Specifically, among the least and most market-oriented firms—and as each firm became more market-oriented—more market-oriented firms exhibited much stronger organizational cultures. Characteristics indicative of a strong culture include organizationwide adherence to values and norms, homogeneity of language and meaning, elaborate methods for selecting and indoctrinating new organization members, dispersion of power among organization members, distribution of rewards and recognition based on behaviors consistent with the culture, and the organization as the primary group identification of members (e.g., O'Reilly and Chatman 1996; Schein 1985). Thus, the market-oriented behaviors we observe are the natural result of deeply held, shared cultural values. These cultural values, however, appear to be necessary, but not sufficient, for the presence of market-oriented behaviors.

Intraorganizational Power

Our research demonstrates that market-oriented organizations are characterized by a particular distribution of intraorganizational power. Power in organizations has been studied extensively (e.g., Pfeffer 1981; Van de Ven and Poole

1995). This work considers an organization a collection of individuals or coalitions, both advancing their own agenda with differing degrees of success in achieving their desired outcomes. In marketing, analyses of interorganizational power play a central role in the study of distribution systems (e.g., Stern and Reve 1980). However, research on market orientation has not incorporated the notion of power. Therefore, it implicitly assumes that all individuals within the organization share a common goal (e.g., Day 1999; Jaworski and Kohli 1993).

In contrast, our research finds that interpersonal and intraorganizational power play a significant role in the change to a market orientation, particularly with regard to instigating change, guiding change, and distributing the use of individual power to maintain a market orientation. Instigation of change to a market orientation begins with a market-oriented coalition gaining control through a dialectical struggle of ideas and power. Having won this struggle, guiding coalitions wield their power to guide cultural change through value inculcation, transformative market experiences, hiring of fellow believers, and removal of cultural dissenters. As organizations institutionalize a market-oriented culture, power devolves from guiding-coalition members to a more egalitarian distribution among organization members. This change from concentrated to distributed power is possible because of the growing strength of a market-oriented culture, and it is also necessary for acting collaboratively in a market-oriented manner. Thus, our research illustrates that the distribution of intraorganizational power is an essential, but overlooked, dimension of market orientation.

Organizational Learning and Schema Development

Organizational learning plays an important, if underappreciated, role in the creation of a market orientation. As it is typically conceived, organizational learning is the collection, interpretation, and encoding of the organization's experience (e.g., Huber 1991; Nonaka 1994). For example, when launching a new marketing strategy, an organization will perceive some outcomes, interpret the causes of those results, and then encode those lessons in the organization through the creation, modification, or elimination of organizational rules, processes, or principles. Analyses of market orientation have typically assumed that market orientation and organizational learning are independent (e.g., Baker and Sinkula 1999; Slater and Narver 1995). Although market-oriented firms may indeed be learning organizations, organizational learning is not considered necessary for a market orientation.

In contrast, our research implies that market-oriented firms fundamentally *are* learning organizations. Through time, members of the organization share common experiences, and those common experiences become formalized as organizationally shared market and process schemas. These schemas enable organization members to communicate and collaborate effectively in the process of gathering, disseminating, and reacting to market intelligence. Creating and using shared schemas is central to organizational researchers' conceptualization of learning organizations (Huber 1991; Nonaka 1994). It is with these shared schemas that organizations can effectively gather, disseminate, and act on explicit knowledge (Nonaka 1994).

The importance of organizational learning for market orientation extends beyond simply encoding the lessons of history. As a result of creating a market orientation, organizations also develop the capacity to evolve. They do so by creating processes for monitoring and changing schemas; this is referred to as a "third-order" change (Bartunek and Moch 1987). The creation of such process schemas is essential for a firm to maintain a market orientation in dynamic markets. By continually verifying and updating market schemas over time through shared experiences, market-oriented firms gain more experience with a market-oriented culture, and by operating in a market-oriented manner, the culture continues to strengthen while becoming increasingly adept at monitoring and reacting to market changes. We believe that this capability is an integral characteristic for maintaining a market orientation. Organizational learning and the schemas that result—market and process schemas that firms use to adapt to the market—are important, yet often-overlooked aspects of a market orientation revealed by our analysis.

Organizational Change

The process of organizational change that we describe differs in important ways from the processes described in the organizational behavior and sociology literature. Within the organizational behavior literature, scholars have identified four principal motors of change: (1) life cycle—predefined steps in an organizational life cycle, (2) evolutionary—evolution through experimentation with new methods, (3) dialectical—political coalitions creating and frustrating change, and (4) teleological—management actions (Van de Ven and Poole 1995). At least 12 major research streams have been identified that incorporate one, two, or three of these change motors (Van de Ven and Poole 1995). Sociologists have identified other modes of change, such as coup d'etat, insurgency, and mass mobilization (Zald and Berger 1978). This literature describes these mechanisms and modes of change in great detail but does not offer any insight into how they might combine to enable an organization to realize a greater market orientation.

Our research illustrates that creating a market-focused organization requires several of these change mechanisms or modes in a specific sequence. The change process begins with the formation of an elite group of insurgents who mobilize the masses to create a consensus for organizational change that is later formalized and sustained. This process suggests a unique combination of change mechanisms and motors. It begins with an insurgency, moves to mass mobilization, switches to a teleological change process, and continues with an indoctrination process to sustain the orientation. Although similar to previous descriptions—each of these stages is recognizable on the basis of similarities to other situations—the combination and sequence is unique.

Moreover, each stage has unique characteristics that have not been previously identified. Consider the insurgents who start the process. The sociological literature discusses an insurgency as a group of disempowered members of an organization who seek to topple the empowered. In contrast, we observed an elite insurgency, an insurgency of the powerful. Likewise, mass mobilization is typically used by disempowered individuals or groups to overwhelm the powerful. In our analysis, however, the insurgents use their power to create a greater enthusiasm for change by mobilizing the masses within the organization. Contrary to the conventional stereotype, they are "elitist revolutionaries" who plot to transform the culture and to create a new order that devolves power to the masses.

Conclusion

Market orientation is a centrally important idea to marketing and a growing number of fields. Although the concept of market orientation has received considerable attention, how organizations develop a greater market orientation has received little attention. Our analysis of four firms successfully making the transformation shows that organizations create a market orientation by engaging in a four-stage process of cultural transformation: (1) initiation, (2) reconstitution, (3) institutionalization, and (4) maintenance. This process imbues the organization with a set of cultural values that support market-oriented activities, an organizationally shared understanding of the market, and organizational learning capabilities. Our results show that these cultural values are central to the new organization that emerges from this process and that the intraorganizational distribution of power and organizational learning play central, though overlooked, roles in creating and sustaining a market orientation.

REFERENCES

Allen, Chris T., Edward F. McQuarrie, and Terri Feldman Barr (1998), "Implementing the Marketing Concept One Employee at a Time: Pinpointing Beliefs About Customer Focus as a Lever for Organizational Renewal," *Journal of Market-Focused Management*, 3 (2), 151–70.

Argyris, Chris (1990), *Overcoming Organizational Defenses: Facilitating Organizational Learning*. Boston: Allyn & Bacon.

Arnould, Eric J. and Craig J. Thompson (2005), "Consumer Culture Theory (CCT): Twenty Years of Research," *Journal of Consumer Research*, 31 (4), 868–82.

——— and Melanie Wallendorf (1994), "Market-Oriented Ethnography: Interpretation Building and Marketing Strategy Formulation," *Journal of Marketing Research*, 31 (November), 484–504.

Baker, William E. and James M. Sinkula (1999), "The Synergistic Effect of Market Orientation and Learning Orientation on Organizational Performance," *Journal of the Academy of Marketing Science*, 27 (4), 411–27.

Bartunek, Jean M. and Michael K. Moch (1987), "First-Order, Second-Order, and Third-Order Change and Organization Development Interventions: A Cognitive Approach," *Journal of Applied Behavioral Science*, 23 (4), 24–41.

Belk, Russell W., John F. Sherry Jr., and Melanie Wallendorf (1988), "A Naturalistic Inquiry into Buyer and Seller Behavior at a Swap Meet," *Journal of Consumer Research*, 14 (4), 449–70.

Besanko, David, David Dranove, and Mark Shanley (2000), *Economics of Strategy*, 2d ed. New York: John Wiley & Sons.

Bonoma, Thomas V. (1985), "Case Research in Marketing: Opportunities, Problems, and a Process," *Journal of Marketing Research*, 22 (May), 199–208.

Burgelman, Robert A. (1983), "A Process Model of Internal Corporate Venturing in the Diversified Major Firm," *Administrative Science Quarterly*, 28 (2), 223–44.

Collins, James C. and Jerry I. Porras (1994), *Built to Last: Successful Habits of Visionary Companies*. New York: Harper Business.

Day, George S. (1999), *The Market Driven Organization: Understanding, Attracting, and Keeping Valuable Customers* New York: The Free Press.

Deshpandé, Rohit (1983), "'Paradigms Lost': On Theory and Method in Research in Marketing," *Journal of Marketing*, 47 (October), 101–110.

——— and John U. Farley (1998), "The Market Orientation Construct: Correlations, Culture, and Comprehensiveness," *Journal of Market-Focused Management*, 2 (3), 237–39.

——— and Frederick E. Webster Jr. (1989), "Organizational Culture and Marketing: Defining the Research Agenda," *Journal of Marketing*, 53 (January), 3–15.

Edwards, Cliff (2005), "Shaking Up Intel's Insides," *Business-Week*, (January 31), 35.

Eisenhardt, Kathleen M. (1989), "Building Theories from Case Study Research," *Academy of Management Review*, 14 (4), 532–50.

Fournier, Susan, James McAlexander, John Schouten, and Sylvia Sensiper (2000), *Building Brand Community on the Harley-Davidson Posse Ride*, rev ed. Boston: Harvard Business School Press.

French, Wendell L. and Cecil Bell (1999), *Organization Development: Behavioral Science Interventions for Organization Improvement*, 6th ed. Upper Saddle River, NJ: Prentice Hall.

Galinsky, Adam D. and Gordon B. Moskowitz (2000), "Perspective-Taking: Decreasing Stereotype Expression, Stereotype Accessibility, and In-Group Favoritism," *Journal of Personality and Social Psychology*, 78 (4), 708–724.

Glaser, Barney G. and Anselm L. Strauss (1967), *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Hawthorne, NY: Aldine de Gruyter.

Gouillart, Francis J. and Frederick D. Sturdivant (1994), "Spend a Day in the Life of Your Customers," *Harvard Business Review*, 72 (1), 116–25.

Homburg, Christian and Christian Pflesser (2000), "A Multiple-Layer Model of Market-Oriented Organizational Culture: Measurement Issues and Performance Outcomes," *Journal of Marketing Research*, 37 (November), 449–62.

Huber, George P. (1991), "Organizational Learning: The Contributing Processes and the Literatures," *Organization Science*, 2 (1), 88–115.

Hult, G. Tomas M., David J. Ketchen Jr., and Stanley F. Slater (2004), "Information Processing, Knowledge Development, and Strategic Supply Chain Performance," *Academy of Management Journal*, 47 (2), 241–53.

Jaworski, Bernard J. and Ajay K. Kohli (1993), "Market Orientation: Antecedents and Consequences," *Journal of Marketing*, 57 (July), 53–70.

Kirca, Ahmet H., Satish Jayachandran, and William O. Bearden (2005), "Market Orientation: A Meta-Analytic Review and

Assessment of Its Antecedents and Impact on Performance," *Journal of Marketing*, 69 (April), 24–41.

Kohli, Ajay K. and Bernard J. Jaworski (1990), "Market Orientation: The Construct, Research Propositions, and Managerial Implications," *Journal of Marketing*, 54 (April), 1–18.

———, ———, and Ajith Kumar (1993), "MARKOR: A Measure of Market Orientation," *Journal of Marketing Research*, 30 (November), 467–77.

Kotler, Philip (2000), *Marketing Management*, millennium ed. Upper Saddle River, NJ: Prentice Hall.

Lincoln, Yvonna S. and Egon G. Guba (1985), *Naturalistic Inquiry*. Beverly Hills, CA: Sage Publications.

Moorman, Christine (1995), "Organizational Market Information Processes: Cultural Antecedents and New Product Outcomes," *Journal of Marketing Research*, 32 (August), 318–35.

Narver, John C. and Stanley F. Slater (1990), "The Effect of a Market Orientation on Business Profitability," *Journal of Marketing*, 54 (October), 20–35.

——— and ——— (1998), "Additional Thoughts on the Measurement of Market Orientation: A Comment on Deshpandé and Farley," *Journal of Market-Focused Management*, 2 (3), 233–36.

———, ———, and Brian Tietje (1998), "Creating a Market Orientation," *Journal of Market-Focused Management*, 2 (3), 241–56.

Nonaka, Ikujiro (1994), "A Dynamic Theory of Organizational Knowledge Creation," *Organization Science*, 5 (1), 14–37.

O'Reilly, Charles A. and Jennifer A. Chatman (1996), "Culture as Social Control: Corporations, Cults, and Commitment," in *Research in Organizational Behavior: An Annual Series of Analytical Essays and Critical Reviews*, Vol. 18, Barry M. Staw and L.L. Cummings, eds. Stamford, CT: JAI Press.

Pettigrew, Andrew M. (1979), "On Studying Organizational Cultures," *Administrative Science Quarterly*, 24 (4), 570–81.

——— (1990), "Longitudinal Field Research on Change: Theory and Practice," *Organization Science*, 1 (3), 267–92.

Pfeffer, Jeffrey (1981), *Power in Organizations*. Marshfield, MA: Pitman.

Reid, Peter C. (1990), *Well Made in America: Lessons from Harley-Davidson on Being the Best*. New York: McGraw-Hill.

Schein, Edgar H. (1985), *Organizational Culture and Leadership*, 1st ed. San Francisco: Jossey-Bass.

Shapiro, Benson P. (1988), "What the Hell Is 'Market Oriented'?" *Harvard Business Review*, 66 (6), 119–25.

Sherry, John F. (1995), *Contemporary Marketing and Consumer Behavior: An Anthropological Sourcebook*. Thousand Oaks, CA: Sage Publications.

———, ed. (1998), *Servicescapes: The Concept of Place in Contemporary Markets*. Lincolnwood, IL: NTC Business Books.

——— and Robert V. Kozinets (2001), "Qualitative Inquiry in Marketing and Consumer Research," in *Kellogg on Marketing*, Dawn Iacobucci, ed. New York: John Wiley & Sons.

Slater, Stanley F. and John C. Narver (1995), "Market Orientation and the Learning Organization," *Journal of Marketing*, 59 (July), 63–74.

Stern, Louis W. and Torger Reve (1980), "Distribution Channels as Political Economies: A Framework for Comparative Analysis," *Journal of Marketing*, 44 (July), 52–64.

Strauss, Anselm L. and Juliet M. Corbin (1998), *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2d ed. Thousand Oaks, CA: Sage Publications.

Teerlink, Rich and Lee Ozley (2000), *More Than a Motorcycle: The Leadership Journey at Harley-Davidson*. Boston: Harvard Business School Press.

Thompson, Craig J., William B. Locander, and Howard R. Pollio (1989), "Putting Consumer Experience Back into Consumer Research: The Philosophy and Method of Existential-Phenomenology," *Journal of Consumer Research*, 16 (2), 133–46.

Turner, Victor Witter (1975), *Dramas, Fields, and Metaphors: Symbolic Action in Human Society*. Ithaca, NY: Cornell University Press.

Van de Ven, Andrew H. and Marshall Scott Poole (1995), "Explaining Development and Change in Organizations," *Academy of Management Review*, 20 (3), 510–40.

Whyte, William F. (1978), "Organizational Behavior Research: Where Do We Go from Here?" in *Applied Anthropology in America: Past Contributions and Future Directions*, Elizabeth M. Eddy and William L. Partridge, eds. New York: Columbia University Press.

Zald, Mayer N. and Michael A. Berger (1978), "Social Movements in Organizations: Coup D'etat, Insurgency, and Mass Movements," *American Journal of Sociology*, 83 (4), 823–61.

Joann Peck & Jennifer Wiggins

It Just Feels Good: Customers' Affective Response to Touch and Its Influence on Persuasion

Prior research has assumed that touch has a persuasive effect only if it provides attribute or structural information about a product. Under this view, the role of touch as a persuasive tool is limited. The main purpose of this research is to investigate the persuasive influence of touch as an affective tool in the *absence* of useful product-related information. The authors find that for people who are motivated to touch because it is fun or interesting, a communication that incorporates touch leads to increased affective response and increased persuasion, particularly when the touch provides neutral or positive sensory feedback. People who are not motivated to touch for fun will also be persuaded by a communication that incorporates touch when they are able to make sense of how the touch is related to the message. The authors explore the effectiveness of different types of touch in generating an affective response, and they replicate the effects on attitudes and behavior in a real-world setting. This research suggests that the marketing implications of touch are more substantial than previously believed. The authors present research implications for direct marketing, product packaging, point-of-purchase displays, and print advertising.

The opportunity to touch products has been shown to have a persuasive influence on customers' attitudes and behavior. Touching a product has been found to increase attitudes and purchase intentions toward the product and to increase the confidence in the evaluation of these products (Peck and Childers 2003a). The need to touch in product evaluation has been linked to the placement of products in stores (Underhill 1999) and to the inability of certain products to be sold online (Citrin et al. 2003; McCabe and Nowlis 2003).

Most applications in marketing focus on touch that provides specific attribute information about the product. For example, the packages of both Paper Mate Dynagrip pens and Ove Glove have portions of the plastic cut out, allowing shoppers to explore and examine the grip of the pen and the unique fabric that makes the gloves heatproof. DuPont created a one-page advertisement for paper used to make overnight courier envelopes that stated, "Go Ahead: Tear this Page in Half" (this appeared in *Fortune Magazine*, December 29, 1997). Attempting the near impossible task of tearing the page conveys the product benefit of the strength of the paper. Such efforts have been shown to have an effect on purchase behavior; for example, providing unwrapped rolls of toilet paper in a point-of-purchase display that allowed customers to feel and compare the textures of different brands resulted in large increases in sales for the store brand in a supermarket chain (Britain's ASDA; see Lindstrom 2005).

For touch to have an influence in marketing decisions and evaluations, must it provide product attribute information, or can hedonic aspects of touch also be persuasive? The importance of hedonic benefits to consumers in marketing has been recognized in the areas of sales promotions (e.g., Chandon, Wansink, and Laurent 2000) and retailing (e.g., Arnold and Reynolds 2003). In this article, we examine whether the hedonic benefits of touch influence decisions independent of the information gathered through touch. We suggest that touch can create an affective response, which can influence a customer's decision-making process even though the touch adds no product-related information to the decision.

If the hedonic aspects of touch can increase persuasion, the use of touch in marketing may be more broadly applicable than previously believed. Thus far, efforts have been limited to touch that provides attribute information about a product; it is often the case that this kind of touch can be used effectively only in contexts in which customers can physically evaluate the product. However, the use of touch as a hedonic tool has the potential to be applied to a broad set of products and even services and in a wide variety of contexts that were previously unrecognized in studies of touch, including package design, print advertising, direct-mail advertising, and point-of-purchase displays.

Joann Peck is Assistant Professor of Marketing, University of Wisconsin-Madison (e-mail: jpeck@bus.wisc.edu). Jennifer Wiggins is Assistant Professor of Marketing, Kent State University (e-mail: jjohnson@bsa3.kent.edu). The authors thank the Madison Children's Museum for its cooperation on this project, and they thank Jan Heide, Ken Wathne, Sung Kim, and Neeraj Arora, as well as colleagues at the University of Wisconsin-Madison, University of Chicago, and the Haring Symposium, for their helpful suggestions and feedback. Finally, they thank the three anonymous *JM* reviewers who helped strengthen this research. This research was funded, in part, by a grant to the first author from the University of Wisconsin-Madison graduate school.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Background and Hypotheses

Instrumental and Autotelic Need for Touch

Peck and Childers (2003b) find that the effects of touch are stronger for some people than for others. They identify individual differences in need for touch (NFT), that is, a person's preference for the extraction and utilization of information obtained through touch. Prior research (see Citrin et al. 2003; Peck and Childers 2003a) has found that some people prefer to evaluate products through touch and are more frustrated when shopping if they do not have the opportunity to touch products. The NFT is conceptualized as having two dimensions: instrumental NFT and autotelic NFT (Peck and Childers 2003b). People who are high in instrumental NFT use touch to gather information about a product to help them make judgments. They are more adept at gathering information through touch, and the opportunity to touch products provides them with access to relevant information they cannot gather through other means, such as reading descriptions of products or visually inspecting products (Peck and Childers 2003a). A customer who is high in instrumental NFT touches a sweater to learn if the material is thick enough to provide warmth. In addition to customers' "rational" information-gathering motives, some people shop for the sensory experiences (Holbrook and Hirschman 1982; Sherry 1990). In contrast, people who are high in autotelic NFT engage in touch because it is fun, interesting, or enjoyable, an experience that is more hedonic than instrumental. A person who is high in autotelic NFT often feels an irresistible need to engage in exploratory touch and is focused on the sensory aspects of touch as an end in and of itself (Peck and Childers 2003b). The customer who fingers the sleeve of a cashmere sweater that he or she has no intention of purchasing, simply because the cashmere feels pleasant to touch, is likely to be high in autotelic NFT.

If touch that does not convey a product attribute can affect persuasion, we expect that this type of touch will be more effective for certain people. Specifically, we posit that a message that incorporates a touch element without product attribute information will be more persuasive for people who are high in autotelic NFT than a message with no touch element. However, we expect that people who are low in autotelic NFT will not find touch as inherently interesting or as irresistible. Consequently, we expect that a touch element that does not provide attribute information will not influence people who are low in autotelic NFT. This leads to our baseline hypothesis:

> H_1: A message that incorporates a touch element (versus a message with no touch element) will increase persuasion for people who are high in autotelic NFT but will not influence persuasion for those who are low in autotelic NFT.[1]

Autotelic NFT and the Sensory Feedback of the Touch Information

It is probable that the type and valence of sensory feedback provided by the touch element influence its persuasiveness. Touch that produces positive sensory feedback has been shown to increase attitudes more for high-NFT people than for low-NFT people. In one study, high-NFT people exhibited a larger increase in attitudes when they touched a soft sweater that was pleasant to touch than when they touched a rough sweater that was less pleasant to touch (Peck 1999). The pleasant sensory experience of touching the soft sweater appears to have increased persuasion more for the high- than for the low-NFT participants. The affective nature of autotelic NFT suggests that people who are high in autotelic NFT are likely to have a stronger affective response to touch than those who are low in autotelic NFT. This implies that people who are high in autotelic NFT, in particular, are likely to be more susceptible to the increase in persuasion that comes from a pleasant touch experience.

Positive affective responses have been shown to influence attitudes and behavior. In advertising research, inducing a positive mood in viewers or generating a positive affective response has been found to increase attitude toward the ad (see Aaker, Stayman, and Hagerty 1986; Batra and Ray 1986; Brown, Homer, and Inman 1998; Burke and Edell 1989), time spent viewing the ad (Olney, Holbrook, and Batra 1991), and attitude toward the brand (Brown, Homer, and Inman 1998; Burke and Edell 1989; Holbrook and Batra 1987). Positive affective responses have also been shown to influence behavior directly. For example, experiencing pleasure (Cunningham 1979; Forbes and TeVault 1975; Isen and Levin 1972; Strahilevitz and Myers 1998) has been shown to increase people's likelihood of donating to charity significantly. Positive feelings have also been shown to increase people's willingness to participate in an experiment and to help people in need (Isen 1987). Finally, positive affect has been linked to the hedonic and experiential aspects of consumer behavior, increasing variety-seeking behavior, experiential shopping, and hedonic consumption (Cohen and Areni 1991; Hirschman and Stern 1999; Kahn and Isen 1993). This suggests that engaging in touch that creates a positive affective response is likely to lead to more positive attitudes and greater behavioral intentions toward a product.

Therefore, we expect that people who are high in autotelic NFT will exhibit an increase in persuasion when they are exposed to marketing messages that incorporate a touch element with positive sensory feedback. In contrast, people who are low in autotelic NFT are more likely to process touch information that is included in a marketing message in the same way that they process any other information in the message and, therefore, will be unlikely to experience a persuasive effect of positive sensory feedback. Thus, touch should influence people who are low in autotelic NFT only if it provides information that helps them interpret the message. Exposure to a touch element that does not provide useful information is unlikely to increase the attitudes of such people. This leads to our next hypothesis:

---

[1] Note that persuasion includes measures of attitude toward the message and attitude toward the organization as well as behavioral measures. For simplicity, we use the term "persuasion."

H_2: Compared with a message with no touch element (control condition), a message that includes a touch element with positive sensory feedback will increase persuasion more for people who are high (but not low) in autotelic NFT than will a message that includes a touch element with neutral or negative sensory feedback.

Congruence Between the Touch Element and the Message

Consumers' responses to additional information that is incorporated into a marketing message, such as a picture or a touch element, have been shown to be influenced by the congruence between the additional information and the message.[2] Heckler and Childers (1992) define "congruence" as having two components: expectancy, or the degree to which the additional information falls into the pattern or structure of the message, and relevancy, or the degree to which the additional information contributes to or detracts from the theme of the message. Information that is unexpected, such as humor, has been shown to increase favorable evaluations and recall of an advertisement (Lee and Mason 1999). Conversely, irrelevant information has been shown to have a negative effect on both the evaluation and the recall of an advertisement (Heckler and Childers 1992; Lee and Mason 1999). This implies that incorporating a touch element into a communication in which it would be unexpected, such as a print advertisement or direct-mail brochure, may lead to a favorable evaluation of the message and increase attitude toward the message, depending on the relevancy of the touch element to the message.

Lee and Mason (1999) suggest that the difficulty with irrelevant information is that a person is unable to perceive the connection between the information and the message and, therefore, becomes frustrated. This implies that if people who are low in autotelic NFT are processing a touch element as part of the overall message, an irrelevant touch element will create confusion and frustration and thus detract from their attitude toward the message. However, if people who are high in autotelic NFT are simply responding affectively to the touch element and not processing it as part of the message, the negative effect of irrelevancy is likely to be attenuated. This leads to our third hypothesis:

H_{3a}: For people who are high in autotelic NFT, a message that includes a touch element will be more persuasive than a message that does not include a touch element, regardless of the congruency between the touch element and the message.

H_{3b}: For people who are low in autotelic NFT, a message that includes a touch element will be more persuasive than a message that does not include a touch element only if the touch element is congruent with the message.

Study 1

The purpose of Study 1 was to test H_1–H_3 and to determine the effects of incorporating a touch element that does not provide product attribute information into a communications message on people who are both high and low in autotelic NFT. Study 1 was a 2 (touch element: present versus absent) × 2 (autotelic NFT: high versus low [determined by a median split]) design; the first factor was manipulated between subjects, and the second factor was measured between subjects. Nested within the touch-element condition were two levels of congruence between the touch element and the message (congruent and incongruent [determined by a pretest]) and three levels of sensory feedback (positive, neutral, and negative [determined by a pretest]); both were manipulated between subjects.

Variables and Procedure

Three hundred forty-five undergraduate students participated in exchange for extra credit in a marketing class. Each participant read a pamphlet with the same message that requested that the participant make a donation of time or money to an arboretum located in the Midwest (see Appendix A). We asked participants to list their thoughts as they read the pamphlet. In the touch-element conditions, a touch element was attached to the front of the pamphlet. We used six touch elements in all; we varied the sensory feedback or valence provided by the touch element and the congruence between the touch element and the message across conditions.

We evaluated congruence of the touch element with the message and sensory feedback through pretests. We presented each participant in the congruency pretest ($n = 56$) with three of the six sample touch elements (we counterbalanced the order and found no order effects) and the pamphlet with the arboretum appeal and asked them to rate the congruence of each sample touch element with the appeal. We measured congruence using two questions: "The 'fit' of this touch sample with the pamphlet is…?" with endpoints "very bad/very good" and "very unfavorable/very favorable," each of which were seven-point scales. We averaged the two measures of congruence (all r's > .85) to obtain one measure of congruence for each touch element. Participants rated three touch elements as highly congruent with the message of the pamphlet: a feather ($M = 5.18$), tree bark ($M = 5.52$), and sandpaper ($M = 4.14$), and none was significantly higher than the others on congruence (*p*s > .05). Participants rated three other touch elements as being low on congruence with the message: a soft silver swatch ($M = 2.42$), a slightly textured black-and-gold swatch ($M = 1.90$), and steel wool ($M = 2.62$), and none was significantly different from the others (all *p*s > .05). The three congruent touch elements were all significantly higher on congruency than the incongruent touch elements.

A second pretest evaluated the valence of the sensory feedback provided by touching each touch element. Thirty-six participants rated how pleasant each of the touch elements was to touch on a seven-point scale. For the congruent touch elements, participants rated the feather as extremely positive ($M = 5.68$), the tree bark as neutral ($M =$

---

[2] The term "congruence" has also been used to describe the relationship between a piece of information that is being processed and a person's information-processing style (see Higgins 2000). Our definition of congruence emphasizes the congruence between the touch information and the message into which it is incorporated and does not imply congruence with the person who is processing the message.

4.01), and the sandpaper as negatively valenced (M = 2.90). Using paired valence t-tests, we obtained the following results: feather versus tree bark: t(35) = 5.84, $p < .01$; feather versus sandpaper: t(35) = 8.45, $p < .01$; and tree bark versus sandpaper: t(35) = 2.42, $p < .01$. Among the incongruent touch elements, participants rated the soft silver swatch as extremely positive (M = 5.59), the slightly textured black-and-gold swatch as neutral (M = 4.04), and the steel wool as negatively valenced (M = 2.90). Again, using paired valence t-tests, we obtained the following results: soft silver swatch versus black-and-gold swatch: t(35) = 6.53, $p < .01$; soft silver swatch versus steel wool: t(35) = 10.60, $p < .01$; and slightly textured black-and-gold swatch versus steel wool: t(35) = 7.80, $p < .01$.[3] After reading the pamphlet, participants in Study 1 completed a questionnaire, which included measures of their attitudes toward the message and the arboretum and their willingness to donate time or money to the arboretum.[4] We used familiarity with the arboretum and prior donation behavior in terms of time and money as covariates in the analyses; we found no significant effects on the results. Next, under the guise of a different study, participants completed the autotelic component of the NFT scale (6 items) from the 12-item NFT scale (Peck and Childers 2003a; for scale items, see Appendix B).[5] The NFT scale measures dimensions of both instrumental NFT and autotelic NFT. As Peck and Childers (2003b) recommend, depending on the underlying theory, researchers might choose to employ just one dimension of the NFT scale. Finally, we thanked participants and debriefed them.

Results

We found support for H_1, which predicted that a touch element would increase persuasion for people who are high in autotelic NFT but not for those who are low in autotelic NFT. We measured both attitude toward the pamphlet and attitude toward the organization on three seven-point scales ("What is your overall feeling toward the pamphlet/organization?" anchored by "very unfavorable/very favorable," "very bad/very good," and "very negative/very positive"). We averaged these (all αs > .90) for a measure of attitude toward the pamphlet and attitude toward the organization. With attitude toward the pamphlet as the dependent variable, the interaction between whether a touch element was present (yes/no) and autotelic NFT was significant (F(1, 339) = 5.25, $p < .05$, $\omega^2 = .02$). There was no significant main effect for the touch/no touch element or for autotelic NFT (*ps* > .05). Using planned contrasts, we found that participants who were high in autotelic NFT had a more positive attitude toward the pamphlet when a touch element was present than when it was absent (Ms = 5.32 and 4.83, respectively; F(1, 339) = 4.93, $p < .05$, $\omega^2 = .04$; see Table 1), whereas a touch element had no effect on participants who were low in autotelic NFT (M = 5.13 for touch and 5.40 for no touch; F(1, 339) = 1.19, $p > .05$). We obtained the same pattern of results for the likelihood of donating time or money to the organization (see Table 1); the interaction between touch/no touch and autotelic NFT was significant (F(1, 339) = 6.16, $p < .05$, $\omega^2 = .02$), but no main effects were significant (*ps* > .05). Participants who were low in autotelic NFT were unaffected by a touch element (Ms = 3.34 for touch and 3.54 for no touch; F(1, 339) = .37, $p > .05$), and participants who were high in autotelic NFT were significantly more willing to donate time or money when a touch element was included in the appeal (Ms = 3.65 and 2.75; F(1, 339) = 9.25, $p < .05$, $\omega^2 = .05$). The presence of the touch element did not affect attitude toward the organization; there were no significant main effects or interaction effects, implying that the persuasive influence of the touch element may be restricted to the specific message that incorporates the touch element and not extended to the organization that is sponsoring the message.

H_2 predicted that for people who are high (but not low) in autotelic NFT, a message that included a touch element with positive sensory feedback (versus a message with no touch element [control condition]) would increase persuasion more than a message that included a touch element with neutral or negative sensory feedback. We found partial support for this hypothesis. Using planned comparisons, for both attitude toward the pamphlet and likelihood of donating time or money, we found that participants who were high in autotelic NFT were most persuaded by touch information that provided positive sensory feedback compared with the no-touch-element control condition (attitude toward the pamphlet: Ms = 5.49 and 4.83; F(1, 329) = 6.32, $p < .05$, $\omega^2 = .05$; likelihood of donating time or money: Ms = 3.87 and 2.75; F(1, 329) = 10.09, $p < .05$, $\omega^2 = .04$; see Table 1). A touch element with neutral sensory feedback also significantly increased persuasion for participants who were high in autotelic NFT compared with the control condition (attitude toward the pamphlet: Ms = 5.32 and 4.83; F(1, 329) = 4.21, $p < .05$, $\omega^2 = .04$; likelihood of donating time or money: Ms = 3.79 and 2.75; F(1, 329) = 9.38, $p < .05$, $\omega^2 = .03$), whereas a touch element providing negative sensory feedback did not significantly influence persuasion for participants who were high in autotelic NFT (all *ps* >

---

[3] We conducted one final pretest to ensure that vision did not influence the touch valence ratings. Sixteen participants evaluated the valence of the six touch elements without vision. We put each touch element in an enclosed box and instructed the participant to reach in and touch the six touch elements, which we presented one at a time with the order counterbalanced. We found no significant order effects, and the valence ratings mirrored those we found in the presence of vision.

[4] Because we used undergraduate students, we also included a measure that stated, "This pamphlet would encourage other people to be more likely to donate time or money to the Arboretum." This question elicited slightly higher means than the students' own willingness to donate, but we obtained the same pattern of results, so we do not report these.

[5] We were concerned that the autotelic dimension of the NFT scale would be too focused on product touch to capture the effects in this context. We administered an additional scale, which included items modified from the autotelic NFT scale to move away from referring to actual product touch (for a list of autotelic NFT items and additional touch items [called "Funtouch"], see Appendix B). We also analyzed all studies using the Funtouch items and found the same results as we did for the autotelic NFT items. For simplicity, we report only the autotelic NFT results.

TABLE 1
Attitude Toward the Pamphlet and Likelihood of Donating Time or Money × Autotelic NFT, Congruence, and Valence: Study 1 Means

	Low Autotelic NFT			High Autotelic NFT		
Valence	**Incongruent**	**Congruent**	**Total**	**Incongruent**	**Congruent**	**Total**
			A: Attitude Toward Pamphlet			
Negative	4.67[a] n = 19	5.06 n = 29	4.90[a] n = 48	5.15 n = 31	5.21 n = 21	5.17 n = 52
Neutral	5.25 n = 24	5.60 n = 21	5.41 n = 45	5.29 n = 25	5.34[b] n = 28	5.32[b] n = 53
Positive	4.74[a] n = 26	5.53 n = 22	5.10 n = 48	5.48[b] n = 21	5.50[b] n = 24	5.49[b] n = 45
Total	4.90[a] n = 69	5.36 n = 72		5.28[b] n = 77	5.36[b] n = 73	
Touch-element total			5.13 n = 141			5.32[b] n = 150
No-touch-element total			5.40 n = 24			4.83 n = 28
			B: Likelihood of Donating Time or Money			
Negative	2.74	3.14	2.98	3.29	3.38	3.33
Neutral	3.04	3.86	3.42	3.88[c]	3.71[c]	3.79[c]
Positive	3.35	3.95	3.62	3.90[c]	3.83[c]	3.87[c]
Total	3.07	3.60		3.65[c]	3.66[c]	
Touch-element total			3.34			3.65[c]
No-touch-element total			3.54			2.75

[a]Cell means are significantly different from the control (no touch element) mean of 5.40 for low autotelic NFT.

[b]Cell means are significantly different from the control (no touch element) mean of 4.83 for high autotelic NFT.

[c]Cell means are significantly different from the control (no touch element) mean of 2.75 for high autotelic NFT.

Notes: All comparisons are based on planned contrasts. All scales are seven-point scales. For likelihood of donating time or money, cell means are significantly different from the control (no touch element) mean of 3.54 for low autotelic NFT.

.05). For participants who were low in autotelic NFT, negative sensory feedback significantly decreased persuasion for attitude toward the pamphlet compared with the no-touch-element control condition (Ms = 4.90 and 5.40; F(1, 329) = 3.94, $p < .05$, $\omega^2 = .02$). However, for likelihood of donating time or money, persuasion was unaffected ($p > .05$; for means, see Table 1). In addition, for participants who were low in autotelic NFT, positive or neutral touch elements did not influence persuasion compared with the no-touch-element condition.

Although not specifically hypothesized, we can examine the effect of valence within the touch-element conditions. Within the touch-element condition, we performed a 3 (valence) × 2 (congruency) × 2 (autotelic NFT) analysis; both attitude toward the pamphlet and likelihood of donating time or money were dependent variables. For both dependent variables, there was a significant main effect for valence of the touch element (attitude toward the pamphlet: F(2, 279) = 3.16, $p < .05$, $\omega^2 = .02$; likelihood of donating time or money: F(2, 279) = 4.84, $p < .05$, $\omega^2 = .04$). There was also a significant main effect of congruence, and a two-way interaction between congruency and autotelic NFT was significant for both dependent variables (we discuss this further with the congruence results); no other interactions were significant. When we used a touch element, a positively valenced or neutral valenced touch element was more persuasive than a touch element that provided negative sensory feedback. For attitude toward the pamphlet, a neutral or positively valenced touch element resulted in a significantly more positive attitude toward the pamphlet than a negatively valenced touch element (negative [M = 5.04] versus positive [M = 5.29] touch element: F(2, 279) = 2.85, $p < .05$, $\omega^2 = .02$; negative versus neutral [M = 5.36] touch element: F(2, 279) = 3.65, $p < .05$, $\omega^2 = .02$). The difference between the mean attitude toward the pamphlet for a neutral touch feedback element and the mean attitude toward the pamphlet for a positive touch feedback element was not significant ($p > .05$).

With likelihood of donating time or money as the dependent variable, we found parallel results. A neutral or positive feedback from the touch element resulted in a significantly greater likelihood of donating time or money to the organization than a touch element that provided negative feedback (negative [M = 3.16] versus positive [M = 3.74] touch element: F(2, 279) = 6.23, $p < .01$, $\omega^2 = .04$; negative versus neutral [M = 3.62] touch element: F(2, 279) = 4.84, $p < .01$, $\omega^2 = .04$). As with attitude toward the pamphlet, the difference between the mean likelihood of donating time or money to the organization with a touch element that provided neutral feedback and a touch element that provided positive feedback was not significant ($p > .05$).

In addition to valence, we examined the effect of the congruency between the touch element and the message. We expected that for participants who were high in autotelic NFT, the congruency between the touch element and the message would not matter; both would be more persuasive than the no-touch-element control (H_{3a}). We found support for this hypothesis. For both dependent variables, pamphlets with both a congruent and an incongruent touch element were more persuasive than the control condition (control versus congruent: attitude toward the pamphlet: Ms = 4.83 and 5.36; F(1, 329) = 4.61, $p < .05$, $\omega^2 = .02$; likelihood of donating time or money: Ms = 2.75 and 3.66; F(1, 329) = 7.52, $p < .05$, $\omega^2 = .04$; control versus incongruent: attitude toward the pamphlet: Ms = 4.83 and 5.28; F(1, 329) = 3.91, $p < .05$, $\omega^2 = .02$; likelihood of donating time or money: Ms = 2.75 and 3.65; F(1, 329) = 8.46, $p < .05$, $\omega^2 = .04$).

For participants who were low in autotelic NFT, we expected that a message with a congruent touch element would be more persuasive than a message with an incongruent touch element, compared with the no-touch-element condition (H_{3b}). We found partial support for this. An incongruent touch element significantly decreased attitude toward the pamphlet compared with the no-touch-element control condition (Ms = 4.90 and 5.40; F(1, 329) = 4.08, $p < .05$, $\omega^2 = .02$), but it did not significantly influence likelihood of donating time or money (Ms = 3.07 and 3.54; $p > .05$). In addition, the presence of a congruent touch element did not significantly influence persuasion compared with the control condition (*ps* > .05; for means, see Table 1).

Again, within the touch-element conditions, a 3 (valence) × 2 (congruency) × 2 (autotelic NFT) analysis yielded not only a significant main effect for valence of the touch element (as we previously reported) but also a significant main effect of congruence (attitude toward the pamphlet: F(1, 279) = 5.06, $p < .05$, $\omega^2 = .02$; likelihood of donating time or money: F(1, 279) = 2.65, $p = .10$, $\omega^2 = .01$). Most important, the expected interaction between congruency and autotelic NFT was significant for both dependent variables (attitude toward the pamphlet: F(1, 279) = 3.59, $p < .05$, $\omega^2 = .02$; likelihood of donating time or money: F(1, 279) = 3.66, $p < .05$, $\omega^2 = .02$); no other interactions were significant. In the touch-element condition, comparing congruent and incongruent touch elements yielded different results for participants who were high and those who were low in autotelic NFT. For those who were high in autotelic NFT, the congruency of the touch element to the communication did not influence either attitude toward the pamphlet (Ms = 5.28 and 5.36; F(1, 329) = .06, $p > .05$) or likelihood of donating time or money to the organization (Ms = 3.34 and 3.54; F(1, 329) = .04, $p > .05$). However, participants who were low in autotelic NFT had a more positive attitude toward the pamphlet when the pamphlet included a congruent touch element than when it included an incongruent touch element (Ms = 5.36 and 4.90; F(1, 329) = 7.70, $p < .05$, $\omega^2 = .03$), and they were also more likely to donate time or money to the organization in the congruent-touch-element condition than in the incongruent-touch-element condition (Ms = 3.60 and 3.07; F(1, 329) = 5.98, $p < .05$, $\omega^2 = .02$; see Table 1). It appears that for participants who were high in autotelic NFT, both congruent and incongruent touch elements had a positive effect on both attitude toward the message and likelihood of donating time or money, but for participants who were low in autotelic NFT, a touch element that was incongruent with the message had a potentially negative effect.

A content analysis of the participants' thoughts about the congruent and incongruent touch elements supported the predictions of H_{3a} and H_{3b}. Of the participants, 69 (20%) expressed confusion about the touch element and frustration at being unable to make sense of the connection between the touch element and the message. Of these participants, 52 were in the incongruent-touch-element condition, suggesting that participants indeed had difficulty making sense of an incongruent touch element and became frustrated as a result. Notably, there was no significant difference between participants who were high and those who were low in autotelic NFT in expressing confusion; 34 participants who were high in autotelic NFT and 35 participants who were low in autotelic NFT expressed confusion about the touch element. However, although the participants who were low in autotelic NFT exhibited a decrease in attitude toward the message in the incongruent condition, the participants who were high in autotelic NFT demonstrated no difference in attitude between the congruent and the incongruent conditions, suggesting that though the participants who were high in autotelic NFT were confused by the incongruent touch element, this confusion did not influence their attitudes.

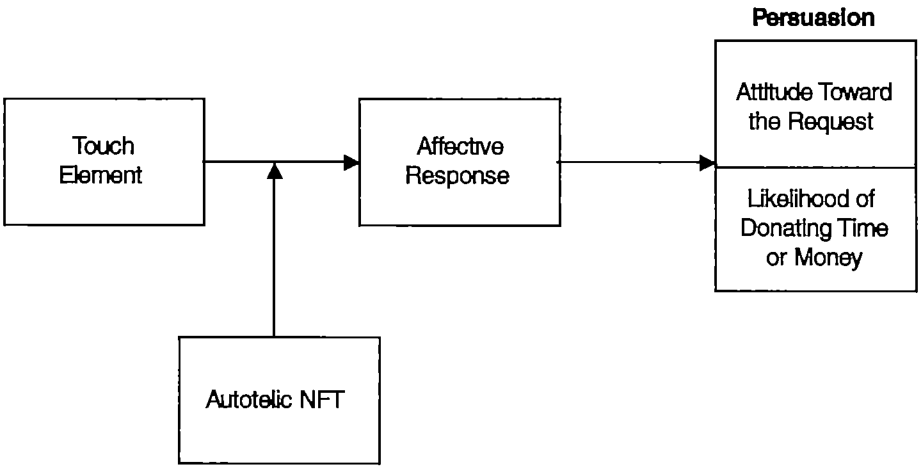
Discussion of Study 1

For participants who were high in autotelic NFT, incorporating a touch element that conveys no product attribute information into a message increased persuasion. This effect occurred regardless of the congruence between the touch element and the message and was stronger for a touch element that provided neutral or positive sensory feedback than for a touch element that provided negative sensory feedback and compared with the no-touch-element control condition. However, participants who were low in autotelic NFT showed no differences in persuasion between the touch and the no-touch-element conditions; they responded negatively to the inclusion of a touch element that was incongruent with the message of the appeal. The results of Study 1 are consistent with the theory that touch creates an affective response for people who are high in autotelic NFT that, in turn, influences persuasion, whereas people who are low in autotelic NFT process touch as part of the message, and touch is persuasive only if it provides useful information. This implies that the relationship among touch, affective response, and persuasion may be a moderated mediation; that is, for people who are high in autotelic NFT, touch creates an affective response that mediates the relationship between touch and persuasion, but for people who are low in autotelic NFT, this mediation effect does not occur (see Figure 1). This leads to our fourth hypothesis:

> H_4: The increased persuasion resulting from a message that incorporates a touch element is mediated by affective response for people who are high in autotelic NFT but not for those who are low in autotelic NFT.

FIGURE 1

Moderated Mediation Affect of Touch on Affective Response and Persuasion


Persuasion
Touch Element
Affective Response
Attitude Toward the Request
Likelihood of Donating Time or Money
Autotelic NFT

Study 2

The purpose of Study 2 was to measure the affective response in both people who were high and those who were low in autotelic NFT to a message that incorporated a touch element and to determine whether this affective response mediated the relationship between touch and persuasion. We also attempted to find additional support for H_1, which predicted that a touch element would be persuasive for people who are high in autotelic NFT but not for those who are low in autotelic NFT. Study 2 was a 2 (autotelic NFT: high versus low [determined by a median split]) × 2 (touch element: present versus absent) design; the first factor was measured between subjects, and the second factor was manipulated between subjects.

Variables and Procedure

Two hundred four undergraduate students participated in the study in exchange for extra credit in a marketing class. Each participant read a pamphlet with the same message that requested that he or she make a donation of time or money to a fictional charity called Spread the Warmth, which provided blankets to needy families during the winter (see Appendix A). In the touch condition, a 4 × 4 inch swatch of black fleece fabric that was similar to the texture of a blanket was attached to the pamphlet; the control condition had no touch element. After reading the pamphlet, participants completed a questionnaire that included measures of their affective response to the pamphlet (for scale items, see Appendix C), attitude toward the pamphlet, attitude toward the organization, and willingness to donate time or money to Spread the Warmth.[6] Next, under the guise of a separate study, we administered the six-item autotelic NFT scale. Finally, we thanked the participants and debriefed them.

Results

With attitude toward the pamphlet as the dependent variable, there was a main effect for whether the touch element was present ($F(1, 200) = 5.25$, $p < .05$, $\omega^2 = .05$), a main effect of autotelic NFT ($F(1, 200) = 13.54$, $p < .01$, $\omega^2 = .08$), and a significant interaction between the two ($F(1, 200) = 4.24$, $p < .05$, $\omega^2 = .04$; for means, see Table 2). Participants who were high in autotelic NFT had a more positive attitude toward the message and a greater likelihood of donating time or money when the pamphlet included a touch element than when it did not (attitude toward the message: Ms = 6.04 and 5.46; $F(1, 200) = 10.39$, $p < .01$, $\omega^2 = .06$; likelihood of donating time or money: Ms = 5.70 and 4.56; $F(1, 200) = 20.75$, $p < .001$, $\omega^2 = .09$), in support of H_1. In contrast, participants who were low in autotelic NFT had no significant difference in attitude toward the message between the touch and the no-touch-element conditions (Ms = 5.27 and 5.24; $F(1, 200) = .02$, $p > .05$). Participants who were low in autotelic NFT showed a significant increase in their likelihood of donating time or money in the touch condition (Ms = 5.00 and 4.36; $F(1, 200) = 5.41$, $p < .05$, $\omega^2 = .03$). For both dependent variables, there was a significant difference between people who were high and those who were low in autotelic NFT in the touch condition, suggesting that those who were high in autotelic NFT responded more strongly to the touch element than those who were low in autotelic NFT (attitude toward the message: Ms = 6.04 and 5.27; $F(1, 200) = 15.67$, $p < .01$, $\omega^2 = .07$; likelihood of donating time or money: Ms = 5.70 and 5.00; $F(1, 200) = 6.78$, $p < .05$, $\omega^2 = .04$; see Table 2). Attitude toward the organization was unaffected by the touch element ($p > .05$).

We conducted Study 2 to examine directly the difference in affective response to the touch element between people who were high and those who were low in autotelic NFT. As expected, we found that participants who were high in autotelic NFT had a stronger affective response to

---

[6] In Study 2, we measured affective response using the two scales in Appendix C. We obtained parallel results for both scales. For simplicity, we report only the results of the first scale. In Study 3, we measured affective response using the emotional reaction scale because we determined that this scale had greater relevance, given the task in which participants were engaged.

TABLE 2

Attitude Toward the Message and Behavioral Intentions × Touch and No-Touch-Element Condtions × High and Low Need for Autotelic Touch: Study 2 Means (Standard Deviations)

	Low Autotelic NFT		High Autotelic NFT	
	No Touch Element	Touch Element	No Touch Element	Touch Element
	n = 55	n = 39	n = 50	n = 60
Attitude toward the pamphlet	5.24	5.27^{b}	5.46^{a}	$6.04^{a, b}$
	(1.10)	(1.02)	(1.01)	(.63)
Attitude toward the organization	6.22	5.94^{a}	6.33	6.35^{a}
	(.80)	(.89)	(.74)	(.84)
Affective response	4.33	4.72^{b}	4.54^{a}	$5.17^{a, b}$
	(1.01)	(1.03)	(1.14)	(1.08)
Likelihood of donating time or money	4.36^{a}	$5.00^{a, c}$	4.56^{b}	$5.70^{b, c}$
	(1.57)	(1.23)	(1.53)	(.79)

Notes: Numbers with the same superscript in the same row are significantly different at $p = .05$, means are based on seven-point scales.

the message in the touch condition than in the no-touch-element condition (Ms = 5.17 and 4.54; F(1, 200) = 9.56, $p < .05$, $\omega^2 = .05$), and those who were low in autotelic NFT had no significant differences in affective response to the message, depending on the presence of a touch element (Ms = 4.72 and 4.33; F(1, 200) = 3.06, $p > .05$; see Table 2). This is consistent with the proposed view that people who are high, but not those who are low, in autotelic NFT are influenced by affective response to touch.

We hypothesized that a touch element that influences persuasion would be mediated by a person's affective response to the touch element. However, we expected autotelic NFT to moderate this relationship (see Figure 1). Using LISREL 8 (Jöreskog and Sörbom 1993), we first estimated a path model using the full sample. For attitude toward the pamphlet, the path model supported complete mediation. The path from the touch element to affective response was significant ($\beta = .23$, t = 3.41, $p < .05$); the path from affective response to attitude toward the pamphlet was significant ($\beta = .50$, t = 8.13, $p < .05$), and the path from the touch element to attitude toward the pamphlet was not significant in the presence of the mediator ($\beta = .08$, t = 1.33, $p > .05$). The results were similar to the dependent variable of likelihood of donating time or money to the organization. The full model showed evidence of partial mediation. The path from the touch element to affective response was significant ($\beta = .23$, t = 3.41, $p < .05$), as was the path from affective response to the likelihood of donating time or money ($\beta = .53$, t = 9.45, $p < .05$). The path from the touch element to donating time or money remained significant in the presence of the mediator, indicating that in addition to mediation, there was a direct effect of the touch element on donating time or money to the organization ($\beta = .22$, t = 3.90, $p < .05$).

To gain further insight, we estimated simultaneous path models for both participants who were high and those who were low in autotelic NFT for each of the two dependent variables. In both cases, partial mediation was supported for those who were high in autotelic NFT, but mediation was not supported for those who were low in autotelic NFT. With attitude toward the pamphlet as the dependent variable, for those who were high in autotelic NFT, the path from the touch element to affective response was significant ($\beta = .26$, t = 2.91, $p < .05$), as was the path from affective response to attitude toward the pamphlet ($\beta = .40$, t = 5.23, $p < .05$). Finally, the path from the touch element to attitude toward the pamphlet remained significant ($\beta = .20$, t = 2.59, $p < .05$), in support of partial mediation. In contrast, for participants who were low in autotelic NFT, the path from the touch element to affective response was not significant ($\beta = .19$, t = 1.40, $p > .05$); thus, mediation of affective response for participants who were low in autotelic NFT was not supported. The path from affective response to attitude toward the pamphlet was significant ($\beta = .81$, t = 5.38, $p < .05$), and there was no significant effect for the touch element on attitude toward the pamphlet ($\beta = -.12$, t = .06, $p > .05$).

The results were parallel for the dependent variable of the likelihood of donating time or money to the organization. For participants who were high in autotelic NFT, the path from the touch element to affective response was significant ($\beta = .21$, t = 3.11, $p < .05$), as was the path from affective response to the likelihood of donating time or money ($\beta = .32$, t = 4.17, $p < .05$). In addition, the path from the touch element to the likelihood of donating time or money was significant ($\beta = .32$, t = 4.17, $p < .05$). However, for participants who were low in autotelic NFT, the path from touch to affective response was not significant ($\beta = .17$, t = 1.56, $p > .05$), in support of differential processes for people who are high and those who are low in autotelic NFT. For participants who were low in autotelic NFT, the path between the touch element and the likelihood of donating time or money was also not significant ($\beta = .13$, t = 1.56, $p > .05$). Not surprisingly, the relationship between affective response and donating time or money to the organization was significant ($\beta = .70$, t = 8.90, $p < .05$).

Discussion of Study 2

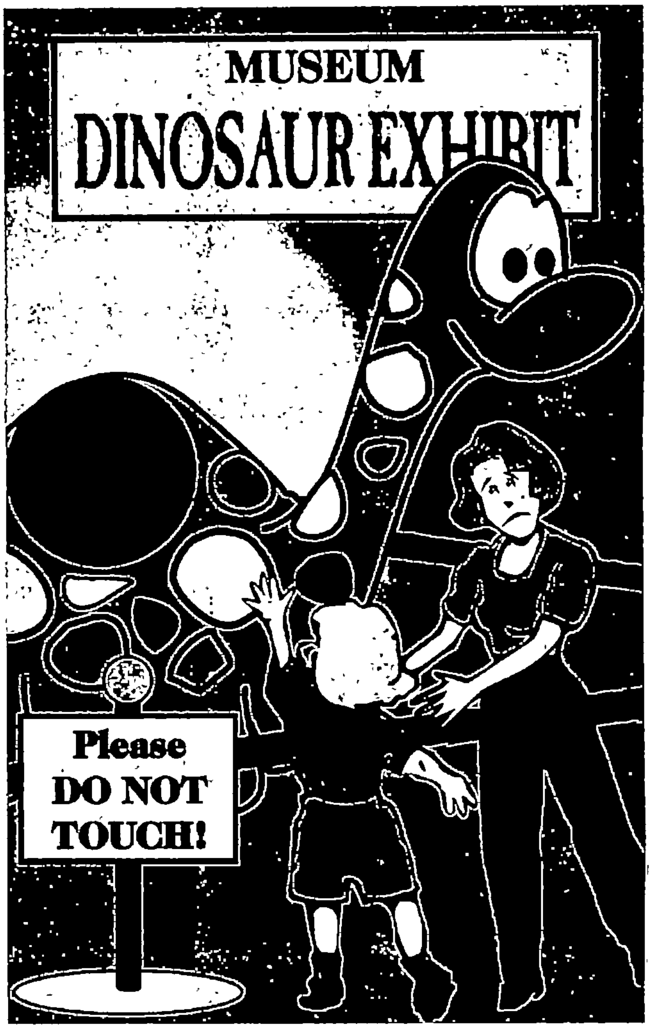
Study 2 replicated the finding that participants who were high in autotelic NFT had a more positive attitude toward messages that incorporated touch elements than toward messages that did not incorporate touch elements. However, this difference did not occur among those who were low in autotelic NFT. We also found that incorporating a touch element increased likelihood of donating time or money among both participants who were high and those who were low in autotelic NFT. This study also directly examined the affective response of those who were high and those who were low in autotelic NFT when encountering a touch element. Only participants who were high in autotelic NFT had a significantly stronger affective response to the message with the touch element than to the message with no touch element. In effect, for participants who were high in autotelic NFT, the touch element evoked an affective response beyond that which the message evoked. We found that this affective response mediated the relationship between touch and persuasion for participants who were high in autotelic NFT but not for those who were low in autotelic NFT.

Study 3

In Study 3, we tested whether the findings of Studies 1 and 2 could be replicated in a real-world situation, and we investigated the generalizability of our experimental findings. We partnered with a midwestern children's museum, a "hands-on" museum targeted to children aged eight and under. As a nonprofit organization, the museum often uses direct-mail appeals to solicit donations and museum memberships. Many of its publicity and fundraising materials stress the museum's unique hands-on features that encourage children to touch, making a touch element relevant to its message. However, the museum had never before incorporated a touch element into one of its direct-marketing efforts. Working with the fundraising staff at the museum, we designed a brochure to solicit new memberships. The cover of the brochure featured a picture of a cuddly looking, spotted, cartoon dinosaur with a child reaching out to touch it (see Figure 2). We then incorporated a touch element into half of the brochures. Because Studies 1 and 2 suggested that touch elements that provide positive sensory feedback are more persuasive, we used a touch element that was soft and pleasant to touch. For people in the touch-element condition, the largest spot on the dinosaur was a circle of red faux fur, which was found in a pretest to be soft and pleasant to touch. For people in the no-touch-element condition, the spot was printed like all the other spots on the dinosaur.

Study 3 was a 2 (autotelic NFT: high versus low [determined by a median split]) × 2 (touch element: present versus absent) design; the first factor was measured between subjects, and the second factor was manipulated between subjects.

FIGURE 2
Brochure for Study 3



Variables and Procedure

We sent the mailing to a purchased list of approximately 2000 families in zip code areas with children under the age of eight who were not currently members of the children's museum. One thousand people on the mailing list received a brochure with a touch element, and the other thousand received the same brochure without a touch element. Each of the 2000 recipients of the brochure also received a questionnaire in the same package, which included questions about the recipient's attitude toward the message and likelihood of donating time or money, control questions about the recipient's prior donations to the museum, and the autotelic NFT scale.[7] We included prior donation behavior and prior visits to the museum as covariates; neither was significant ($p > .05$).

Sample Description

Of the 2000 questionnaires, 116 were returned, for a response rate of 5.8%. According to the Direct Marketing Association (2003), the overall average response rate for direct mail, including mailings to both house and prospect files, is 2.54%, suggesting that our response rate is reasonable. To help eliminate the concern of nonresponse bias, we compared questionnaires returned in the first week (n = 79, or 68%) with those returned in the subsequent three weeks; we found no significant differences in demographics or in independent or dependent variables. Of the responses, 61

---

[7] Because of survey length restrictions the organization imposed, we included only autotelic NFT.

(53%) were from the touch-element condition, and 55 (47%) were from the no-touch-element condition. The majority of our respondents were between the ages of 35 and 44 (53%), followed by ages 25–44 (37%), 45–54 (9%), and 60–64 (1%). The mailing-list company stated that 55% of its list were between the ages of 35 and 44 (compared with our 53%). Female respondents significantly outnumbered male respondents (102 women [88%] and 14 men [12%]).

Results

We expected that the touch element would increase attitude toward the message and behavior for participants who were high but not for those who were low in autotelic NFT. We hoped that we would be able to obtain a measure of actual membership resulting from the mailing. Unfortunately, only 3 of the 116 respondents that returned the survey became members of the museum, so we were unable to use this measure. We measured behavioral intentions through a question that asked, "After reading the brochure, how likely are you to become a member of the Children's Museum?" with endpoints "very unlikely" (1) and "very likely" (7). There was a main effect of the touch element on both attitude toward the pamphlet and likelihood of becoming a member of the museum (attitude toward to pamphlet: Ms = 4.73 and 5.34; no touch element versus touch element: $F(1, 112) = 7.42$, $p < .05$, $\omega^2 = .04$; likelihood of becoming a member: Ms = 3.25 and 3.92; $F(1, 112) = 4.68$, $p < .05$, $\omega^2 = .03$). However, this was qualified by a significant interaction. In support of H_1, the expected interactions between the presence or absence of a touch element and autotelic NFT for both attitude toward the brochure and likelihood of becoming a museum member were significant (attitude toward the brochure: $F(1, 112) = 10.57$, $p < .01$, $\omega^2 = .09$; likelihood of becoming a member: $F(1, 112) = 5.23$, $p < .05$, $\omega^2 = .05$).

Using planned contrasts, we found that participants who were high in autotelic NFT had a more positive attitude toward the pamphlet and would be more likely to become a member of the museum when a touch element was present than when it was absent (attitude toward the pamphlet: Ms = 5.58 and 4.26; $F(1, 112) = 6.44$, $p < .05$, $\omega^2 = .05$; likelihood of becoming a member: Ms = 4.39 and 3.43; $F(1, 112) = 4.01$, $p < .05$, $\omega^2 = .04$; see Table 3). In contrast, as we expected, participants who were low in autotelic NFT were not influenced by the touch element (attitude toward the pamphlet: $F(1, 112) = 1.53$, $p > .05$; likelihood of becoming a member: $F(1, 112) = .55$, $p > .05$). For attitude toward the pamphlet as the dependent variable, we found no main effect of autotelic NFT ($p > .05$); however, we did find a significant main effect of autotleic NFT on likelihood of becoming a member (Ms = 4.03 and 3.13 for participants who were high and those who were low in autotelic NFT, respectively; $F(1, 112) = 6.51$, $p < .05$, $\omega^2 = .06$). As in Studies 1 and 2, the presence of a touch element did not influence attitude toward the organization (all *p*s > .05).

Discussion of Study 3

Study 3 replicated the results of Study 1 in a real-world setting, using a more heterogeneous population than the previous studies. Thus, we are more confident that our results generalize to the general population. A touch element that provided positive sensory feedback incorporated into a real marketing brochure with a congruent message increased attitude toward the brochure and behavioral intentions among recipients who were high in autotelic NFT without diminishing attitude toward the brochure or behavioral intentions among those who were low in autotelic NFT. This demonstrates that incorporating a touch element into a message indeed results in a net increase in persuasion, even if the touch element does not provide any additional instrumental information.

General Discussion

In three studies, we found that the incorporation of touch into marketing messages can have a positive effect on persuasion for people who are high in autotelic NFT. When a touch element was used, a positively valenced or neutral element was more persuasive than when a touch element that provided negative sensory feedback was used. For participants who were high in autotelic NFT, compared with a no-touch-element control condition, a positively valenced or neutral touch element increased persuasion, and a nega-

TABLE 3
Attitude Toward the Message and Behavioral Intentions × Touch and No-Touch-Element Conditions × High and Low Need for Autotelic Touch: Study 3 Means (Standard Deviations)

	Low Autotelic NFT		High Autotelic NFT	
	No Touch Element	Touch Element	No Touch Element	Touch Element
	n = 32	n = 23	n = 23	n = 38
Attitude toward the pamphlet	5.07	5.00	4.26^a	5.58^a
	(1.18)	(1.66)	(.84)	(1.02)
Attitude toward the organization	5.58	5.62	5.96	5.91
	(1.45)	(1.45)	(1.22)	(1.14)
Likelihood of becoming a member	3.12	3.13	3.43^a	4.39^a
	(1.11)	(1.82)	(2.02)	(1.53)

Notes: Numbers with the superscript in the same row are significantly different at $p = .05$; means are based on seven-point scales.

tively valenced touch element did not influence persuasion. However, for participants who were low in autotelic NFT, compared with the no-touch-element condition, a touch element that provided negative feedback decreased persuasion, whereas a touch element that provided positive or neutral feedback did not influence persuasion.

In Study 2, we examined the process by which a touch element that provided positive sensory feedback influenced persuasion. We found that the persuasive effect occurred because of an affective or emotional response to the experience of touch. Participants who were high in autotelic NFT experienced an emotional response due to the touch element, but participants who were low in autotelic NFT did not exhibit this same response. This emotional response to the touch element mediated the relationship between the presence of the touch element and persuasion

In addition to the types of sensory feedback the touch element elicited, we examined the influence of the congruency of the touch element with the overall message. For participants who were high in autotelic NFT, a touch element increased persuasion, regardless of whether the touch element was congruent with the overall marketing message. However, for participants who were low in autotelic NFT, incorporation of a touch element that was not congruent with the message actually decreased the persuasiveness of the message.

Note also that in both Study 2 (the "Feel the Warmth" study) and Study 3 (the children's museum field study), we found a main effect of the presence of a touch element on persuasiveness. This is likely because both of these studies used a touch element that provided positive sensory feedback, which participants perceived as fitting the persuasive message.

Limitations and Further Research

Although this research found strong consistent effects that suggest that touch has a positive affective influence on persuasion, we should recognize some limitations. We were not successful in our attempt to obtain a behavioral measure of persuasion. Further research should examine the effects of touch on actual behavior. In addition, we were unable to find an effect of touch on attitude toward the organization. Although an increase in attitude toward the message and behavioral intentions are positive outcomes for marketers, the focus on building relationships with consumers that has become prevalent in the marketing literature suggests that persuasive elements would be more valuable if they could contribute to a broader attitude toward the organization. Further research may be able to determine whether there are conditions under which the persuasive effects of touch can be extended to the organization that sponsors the message. For example, repeated exposures to the message may influence attitude and behavior toward the organization.

Another possible limitation is that we examined only the process of affective response using a touch element that was positively valenced. In particular, we used softness because it has been associated with a pleasant sensory feeling and has been used in touch research (Bushnell and Boudreau 1991; Essick, James, and McGlone 1999). Further research should continue to explore process issues by examining the affective response to neutral and negatively valenced touch elements. It could also be the case that the emotional response to the touch element is moderated by other factors. In the current research, participants who were high in autotelic NFT exhibited a stronger affective response to touch elements than did participants who were low in autotelic NFT. We know from previous research (Peck and Childers 2003b) that touch information is also more accessible for high- than for low-NFT people. Thus, not only do high-NFT people respond more strongly to touch information, but they also may weigh it more heavily when forming evaluations. Thus, it seems likely that a touch element could be effective for high-NFT people under both high- and low-involvement conditions. Further research should be conducted to examine the issue of involvement as it relates to high- and low-NFT people and the processes of evaluation and persuasion.

Implications

Theoretically, this research extends touch research in marketing. Previous research has focused solely on product touch, which has been found to influence persuasion when it provides instrumental attribute or structural information about the product (e.g., McCabe and Nowlis 2003; Peck and Childers 2003a). Our research found that touch can also be used as a persuasive element outside of the product touch context by providing an enjoyable hedonic experience for the consumer. In addition, this research examined the process by which hedonic touch influences persuasion and the differences between people who are high and those who are low in autotelic NFT. For people who are high in autotelic NFT, incorporating a touch element into a persuasive communications message creates an affective response that increases attitude toward the message and behavioral intentions. For people who are low in autotelic NFT, a touch element does not generate a significant increase in their affective response.

This research implies that touch could be incorporated into marketing messages in a variety of contexts. People who are high in autotelic NFT will experience affective responses and greater persuasion when they receive a message that incorporates touch, especially when the touch element provides neutral or positive sensory feedback. However, to avoid decreasing persuasion for people who are low in autotelic NFT, marketers must ensure that the touch element is congruent with the overall marketing message, and it should not be negatively valenced. That said, the various touch elements we used suggests that congruency between the touch element and the message can be interpreted more broadly than previous research on congruency has suggested. For example, in our study, participants evaluated sandpaper as being congruent with an arboretum. A similar piece of sandpaper was used in an actual direct-mail brochure that a charity used to solicit donations to help children living in poverty in India. On the front of the brochure, there was a two-inch square of sandpaper, and underneath it were the words, "TOUCH THIS...." When the pamphlet was opened, there was a picture of a boy, and underneath the picture were the words, "To Feel 9-year-old Mallesh's hand." The experience of touching the sandpaper likely creates an image of the child's suffering and generates an affective response of sympathy, which may lead people to make donations to the charity. This suggests that marketers can use touch in a variety of contexts, provided that people who are low in autotelic NFT will be able to find some way to make sense of the touch element.

Touch can be used as a persuasive tool beyond brochures that request charitable donations. In some direct-marketing situations, it may be possible to segment customers on the basis of autotelic NFT. Many direct marketers, such as Land's End, currently provide customers with the opportunity to request touch information, such as fabric swatches. It is likely that a person who requests a fabric swatch is high in autotelic NFT because prior research has found a correlation between instrumental and autotelic NFT (e.g., Peck and Childers 2003a, b). Although it may be too expensive to mail a touch-enhanced catalog to all customers, a direct marketer could customize mailings to customers who request fabric swatches.

Regarding the area of product packaging, the current research suggests that the packaging opportunities are not limited to providing touch attribute information. New print technologies are being introduced that provide tactile effects, which encourage consumers to touch (Kaleido 2004). It is likely that people who are high in autotelic NFT are more likely to approach and examine such packages than are people who are low in autotelic NFT, because the former enjoys the sensory experiences of touch. A product package that is interesting to touch may increase sales of the product even if the opportunity to touch does not provide additional product attribute information. Some evidence suggests that tactile elements of product packaging can even contribute to the overall brand image of a product. For example, Lindstrom (2005) discusses Coca-Cola's use of the nostalgic glass bottle to reinforce its brand image and suggests that it is the tactile sensation, the feel of the bottle in the customer's hand, that is associated with the brand.

Note that in this research, we investigated the sensory aspects of touch independent of product attribute information. However, product touch that conveys attribute information may also provide interesting sensory feedback, which would likely affect attitude toward the product beyond the touch attribute information. For example, the courier envelopes we mentioned previously convey the product benefit of the strength of the paper. This research suggests that the smooth, pleasant feel of the paper can also increase product evaluation.

Touch also has significant implications for in-store and point-of-purchase displays. People who are high in autotelic NFT are drawn to opportunities to touch and are likely to respond to opportunities to touch clothing, paper goods, and other products that provide positive sensory feedback, even if they are not in the process of evaluating the product. A display that encourages touch may lead customers to interact with products that they otherwise would have ignored, which in turn may increase impulse and unplanned purchases (Peck and Childers 2006).

Finally, this study suggests that touch can be used along with pictures, photos, color, humor, and other elements to increase the persuasiveness of print advertising. Touch elements are unexpected information and have been shown to increase the persuasiveness of advertisements, provided that they are congruent with the message (Lee and Mason 1999). Recent trends in advertising have focused on the experiential and aesthetic aspects of communication (e.g., Schmitt 1999; Schmitt and Simonson 1997). Incorporating touch may be the next step in adding a hedonic or experiential aspect to advertising and other marketing communications.

Appendix A
Text of Messages for Studies 1 and 2

Study 1

A little piece of nature, nestled in the middle of the city, the Arboretum is a perfect place for a stroll, curling up under a tree with a good book, watching birds and local wildlife, or an evening of stargazing.

But the Arboretum is more than just a park. It is a research and teaching facility that provides a place for people to develop a positive relationship with nature.

Every day, the Arboretum brings more people back in touch with nature with environmental tours, class and lectures, and special outreach programs at local schools. It is an exceptional resource for learning, sharing, discovering, and enjoying—a cherished gem of nature in an urban setting.

The Arboretum also provides great opportunities to get in touch with nature directly with hiking, biking and jogging trails, and even trails for skiing and snowshoeing.

The Arboretum has been a pioneer in the restoration, and management of ecological communities since the 1930's. The Arboretum strives to conserve, restore and preserve the natural lands of the city so that residents can enjoy the beauty of nature for years to come.

As part of this goal, we recently embarked on a capital campaign to further improve the Arboretum through a new addition to the visitor's center and the installation of the four-acre Native Plant Garden. This project will help the Arboretum improve and expand programs for university, public and professional audiences.

You can help to bring more people in touch with nature by becoming a Friend of the Arboretum. We gratefully welcome your gift because every gift enhances our ability to provide quality programs and experiences for everyone who enjoys and learns from the Arboretum.

Study 2

Feel the warmth of a warm winter blanket.

Winters can be icy cold, especially at night. There's nothing like curling up under a thick, cozy, warm blanket to keep out the harsh winter chill. But for some families, that warm, cozy feeling is just out of reach.

Spread the Warmth helps families in need to keep warm this winter by providing them with new and gently used blankets. But we need your help to spread the warmth before the first big winter chill hits our state this year.

Spread the Warmth needs volunteers to deliver blankets in our city and donations to help us to purchase enough blankets to keep everyone warm. Please help us spread the warmth this winter.

Appendix B
Autotelic NFT Items

Autotelic NFT (6 Items)

1. When walking through stores, I can't help touching all kinds of products.
2. Touching products can be fun.
3. When browsing in stores, it is important for me to handle all kinds of products.
4. I like to touch products even if I have no intention of buying them.
5. When browsing in stores, I like to touch lots of products.
6. I find myself touching all kinds of products in stores.

We measured all items on a seven-point Likert-type scale: $\alpha = .92$ (source: Peck and Childers 2003b).

Funtouch (3 Items)

1. I enjoy touching various textures.
2. I am a person who likes to touch.
3. Touching in general is fun.

We measured all items on a seven-point Likert-type scale: $\alpha = .94$.

Appendix C
Affective Response Items

Affective Response (4 items)

1. This mailing was very enjoyable.
2. This mailing was very likeable.
3. This mailing was very persuasive.
4. This mailing was very interesting.

We measured all items on a seven-point Likert-type scale: $\alpha = .86$ (source: Zinkhan and Martin 1983).

Emotional Reaction (10 Items)

Here is a list of emotional reactions you may have experienced while reading the mailing (touching the swatches). Please indicate how much you felt each of these emotional reactions.

- Interested
- Moved
- Captivated
- Inquiring
- Confident
- Delighted
- Enthusiastic
- Appealed
- Satisfied
- Amused

We measured all items on a five-point scale, with endpoints "not at all" and "a lot": $\alpha = .84$ (source: Derbaix 1995).

REFERENCES

Aaker, David A., Douglas M. Stayman, and Michael R. Hagerty (1986), "Warmth in Advertising: Measurement, Impact, and Sequence Effects," *Journal of Consumer Research*, 12 (March), 365–81.

Arnold, Mark J. and Kristy E. Reynolds (2003), "Hedonic Shopping Motivations," *Journal of Retailing*, 79 (Summer), 77–95.

Batra, Rajeev and Michael L. Ray (1986), "Affective Responses Mediating Acceptance of Advertising," *Journal of Consumer Research*, 13 (September), 234–49.

Brown, Steven P., Pamela M. Homer, and J. Jeffrey Inman (1998), "A Meta-Analysis of Relationships Between Ad-Evoked Feelings and Advertising Responses," *Journal of Marketing Research*, 35 (February), 114–26.

Burke, Marian Chapman and Julie A. Edell (1989), "The Impact of Feelings on Ad-Based Affect and Cognition," *Journal of Marketing Research*, 26 (February), 69–83.

Bushnell, Emily W. and J. Paul Boudreau (1991), "The Development of Haptic Perception During Infancy," in *The Psychology of Touch*, Morton A. Heller and William Schiff, eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 139–61.

Chandon, Pierre, Brian Wansink, and Giles Laurent (2000), "A Benefit Congruency Framework of Sales Promotion Effectiveness," *Journal of Marketing*, 64 (October), 65–81.

Citrin, Alka Varma, Donald E. Stem, Eric R. Spangenberg, and Michael J. Clark (2003), "Consumer Need for Tactile Input: An Internet Retailing Challenge," *Journal of Business Research*, 56 (11), 915–22.

Cohen, Joel B. and Charles S. Areni (1991), "Affect and Consumer Behavior," in *Handbook of Consumer Behavior*, Thomas R. Robertson and Harold H. Kassarjian, eds. Englewood Cliffs, NJ: Prentice Hall, 188–240.

Cunningham, Michael R. (1979), "Weather, Mood and Helping Behavior: Quasi Experiments with the Sunshine Samaritan," *Journal of Personality and Social Psychology*, 37 (November), 1947–56.

Derbaix, Christian (1995), "The Impact of Affective Reactions on Attitudes Toward the Advertisement and the Brand: A Step Toward Ecological Validity," *Journal of Marketing Research*, 32 (November), 470–79.

Direct Marketing Association (2003), "The DMA's 2003 Response Rate Study," (accessed October 23, 2003), [available at http://www.the-dma.org/].

Essick, Greg Anders, Anuj James, and Francis P. McGlone (1999), "Psycholophysical Assessment of the Affective Components of Non-Painful Touch," *Neuroreport: For Rapid Communication of Neuroscience Research*, 10 (10), 2083–2087.

Forbes, Gordon B. and Richard K. TeVault (1975), "The Facilitation of Anonymous Helpfulness by a Fortuitous Pleasant Event," *Journal of Social Psychology*, 97 (December), 299–300.

Heckler, Susan E. and Terry L. Childers (1992), "The Role of Expectancy and Relevancy in Memory for Verbal and Visual Information. What Is Incongruency?" *Journal of Consumer Research*, 18 (March), 475–92.

Higgins, E. Tory (2000), "Making a Good Decision: Values from Fit," *American Psychologist*, 55 (November), 1217–30.

Hirschman, Elizabeth C. and B.B. Stern (1999), "The Roles of Emotion in Consumer Research," in *Advances in Consumer Research*, Vol. 26, Eric J. Arnould and Linda M. Scott, eds. Provo, UT: Association for Consumer Research, 4–11.

Holbrook, Morris B. and Rajeev Batra (1987), "Assessing the Role of Emotions as Mediators of Consumer Responses to Advertising," *Journal of Consumer Research*, 14 (December), 404–420.

——— and Elizabeth C. Hirschman (1982), "The Experiential Aspects of Consumption: Consumer Fantasies, Feelings, and Fun," *Journal of Consumer Research*, 9 (September), 132–40.

Isen, Alice M. (1987), "Positive Affect, Cognitive Processes, and Social Behavior," in *Advances in Experimental Social Psychology*, Vol. 20, L. Berkowitz, ed. New York: Academic Press, 203–53.

——— and Paula F. Levin (1972), "The Effect of Feeling Good on Helping: Cookies and Kindness," *Journal of Personality and Social Psychology*, 21 (March), 384–88.

Jöreskog, Karl G. and Dag Sörbom (1993), *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Chicago: Scientific Software International.

Kahn, Barbara E. and Alice M. Isen (1993), "The Influence of Positive Affect on Variety Seeking Among Safe, Enjoyable Products," *Journal of Consumer Research*, 20 (2), 257–70.

Kaleido, Leanie (2004), "A Soft Touch for Aerosols," (accessed February 6, 2005), [available at http://www.packagingmagazine.co.uk].

Lee, Yih Hwai and Charlotte Mason (1999), "Responses to Information Incongruency in Advertising: The Role of Expectancy, Relevancy, and Humor," *Journal of Consumer Research*, 26 (September), 156–69.

Lindstrom, Martin (2005), *Brand Sense: Build Powerful Brands Through Touch, Taste, Smell, Sight, and Sound*. New York: The Free Press.

McCabe, Deborah Brown and Stephen M. Nowlis (2003), "The Effect of Examining Actual Products or Product Descriptions on Consumer Preference," *Journal of Consumer Psychology*, 13 (4), 431–39.

Olney, Thomas J., Morris B. Holbrook, and Rajeev Batra (1991), "Consumer Responses to Advertising: The Effects of Ad Content, Emotions, and Attitude Toward the Ad on Viewing Time," *Journal of Consumer Research*, 17 (March), 440–53.

Peck, Joann (1999), "Extraction of Haptic Properties: Individual Characteristics and Stimulus Characteristics," doctoral dissertation, Department of Marketing, University of Minnesota.

——— and Terry L. Childers (2003a), "To Have and to Hold: The Influence of Haptic Information on Product Judgments," *Journal of Marketing*, 67 (April), 35–48.

——— and ——— (2003b), "Individual Differences in Haptic Information Processing: The 'Need for Touch' Scale," *Journal of Consumer Research*, 30 (December), 430–42.

——— and ——— (2006), "If I Touch It I Have to Have It: Effects of Need for Touch on Impulse Purchasing," *Journal of Business Research*, 59 (6), 765–69.

Schmitt, Bernd (1999), *Experiential Marketing: How to Get Customers to Sense, Feel, Think, Act, and Relate to Your Company and Brands*. New York: The Free Press.

——— and Alex Simonson (1997), *Marketing Aesthetics: The Strategic Management of Brands, Identity and Image*. New York: The Free Press.

Sherry, John F., Jr. (1990), "A Sociocultural Analysis of a Midwestern Flea Market," *Journal of Consumer Research*, 17 (June), 13–30.

Strahilevitz, Michal and John G. Myers (1998), "Donations to Charity as Purchase Incentives: How Well They Work May Depend on What You Are Trying to Sell," *Journal of Consumer Research*, 24 (March), 434–46.

Underhill, Paco (1999), *Why We Buy, The Science of Shopping*. New York: Simon and Schuster.

Zinkhan, George M. and Claude R. Martin Jr. (1983), "Message Characteristics and Audience Characteristics: Predictors of Advertising Response," in *Advances in Consumer Research*, Vol. 10, Richard P. Bagozzi and Alice M. Tybout, eds. Provo, UT: Association for Consumer Research, 27–31.

Xueming Luo & Naveen Donthu

Marketing's Credibility: A Longitudinal Investigation of Marketing Communication Productivity and Shareholder Value

Practitioners and scholars have questioned marketing communication's credibility because it has not been linked to firm shareholder value. Using secondary longitudinal data, the authors show that the impact of marketing communication productivity (MCP) on shareholder value is positive and nonlinear. In particular, the influence of MCP on the forward-looking performances of Tobin's q and stock return is curvilinear with an inverted U shape. The authors also investigate the moderating role of research-and-development investment and competitive intensity. The findings have implications for marketing managers regarding how to model and benchmark MCP and the dangers of blindly improving the productivity of advertising and sales promotion expenditures.

Marketing scholars and practitioners have long been concerned about the accountability of marketing communication expenditures (e.g., advertising, sales promotions). In their classic article, Aaker and Carman (1982) report that firms greatly overadvertise. In an echo of this claim, Prasad and Sen (1999), in their review of studies since 1982, find a pattern of continued overadvertising across firms. More recently, Rust and colleagues (2004) comprehensively review the literature and pinpoint various researchers' concerns about the productivity of advertising and sales promotion expenditures. In particular, to rebuild confidence in advertising and marketing investment, Rust and colleagues (2004, p. 76) suggest the urgent need to "show how marketing adds to shareholder value, [because] the perceived lack of accountability has undermined marketing's credibility, threatened marketing's standing in the firm, and even threatened marketing's existence as a distinct capability within the firm."

The trade press has identified a similar problem. According to McLaughlim (1997, p. 5), "the most sacred word in advertising had been creativity; now it's accountability." Clients are demanding the most cost-efficient media planning with measurable returns. Kumar and Petersen (2004, p. 28) point out that "company [chief financial officers] are charged with the task of getting the most out of company resources [such as] marketing communication investments." Kraft Foods has suggested moving away from guesswork and gut feelings and toward the systematic assessment of advertising productivity (Dittus and Kopp 1990). Achenbaum (1992, p. 25) calls attention to the "crisis of advertising productivity" because many customers tend to zap past, zip through, and tape over marketers' multimillion dollar selling and marketing messages. More recently, it has been noted that "there's hardly a marketing executive today who isn't demanding a more scientific approach to help defend marketing strategies from the chief financial officer" (*BusinessWeek* 2004). Thus, reversing the marketing productivity crisis represents a high-profile, hot-button issue for both business practitioners and marketing scholars.

However, several challenges hinder marketers from fully understanding and effectively improving marketing communication accountability. First, an appropriate measurement is needed of marketing communication productivity (MCP), which we define as the ratio of marketing outputs (sales level, sales growth, and corporate reputation) to marketing communication expenditures (broadcast, print, and outdoor advertising and sales promotion expenditures). A methodologically sound measure of MCP remains challenging because firms often target their advertising and sales promotion expenditures to promote multiple outcomes simultaneously, such as both visible sales and market performance and invisible customer equity and corporate brand reputation. Thus, the traditional approach of using a single outcome-based measure (e.g., simple ratio of sales to marketing expenditures) may not offer a valid assessment of MCP. In addition, although a vast literature pertains to sales responses to advertising (e.g., Vakratsas and Ambler 1999), the theorized shape of the response curve (i.e., linear, nonlinear concave, or convex) has been, at best, inconsistent and inconclusive (see Holstius 1990). In this sense, a subjectively specified measure based on the conflicting advertising–sales function that uses regression techniques is

---

Xueming Luo is Assistant Professor of Marketing, Department of Marketing, University of Texas at Arlington (e-mail: luoxm@uta.edu). Naveen Donthu is Katherine S. Bernhardt Research Professor of Marketing, Department of Marketing, Georgia State University (e-mail: ndonthu@gsu.edu). The authors thank the anonymous *JM* reviewers for their helpful guidance and insightful suggestions. Ruth Bolton served as editor for this article.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

inappropriate for assessing MCP. Extant marketing studies have not addressed these challenges sufficiently with rigorous methodological techniques.

Second, there needs to be a proper assessment of the dynamic contributions of MCP to the firm's market value over time. Specifically, in assessing the value of marketing communications, most prior research has related advertising and sales promotions to accounting performance measures, such as profits (e.g., Dekimpe and Hanssens 1995; Mueller 1990). However, Aaker and Jacobson (2001, p. 485) challenge this practice by proposing that accounting-based measures of return on investment (ROI) and book equity alone "cannot adequately explain firm value, because they fail to capture the benefits of investing in intangible assets such as brands." Similarly, Ambler (2004) notes that ROI is driven by short-term considerations and often fails to reflect the dynamic impact of marketing assets. In addition, because accounting measures such as ROI are backward looking in nature (i.e., capture firm value in terms of historical or concurrent performance), they are not robust for evaluating the future contribution of MCP (Rust et al. 2004).[1] Therefore, the long-term value of the invisible outcomes (e.g., corporate reputation, brand equity) of a marketing expenditure is less likely to be evident according to accounting book values. Instead, research should address how MCP adds to a firm's market value in terms of forward-looking stock market performance, such as Tobin's q (Dutta, Narasimhan, and Rajiv 1999; Rao, Agarwal, and Dahlhoff 2004) and stock returns (Aaker and Jacobson 1994; Mizik and Jacobson 2003). Unfortunately, prior literature has rarely related MCP to the market-based measure of firm value and thus has failed to address the real, long-term credibility of marketing communication expenditures.

Against this background, we attempt to (1) develop a theory-based measurement of MCP and (2) assess MCP's association with firm shareholder value empirically. In particular, we define MCP on the basis of economic growth theory and the resource-based view (RBV) and measure it in a time-series-based context using the Malmquist modeling method (Caves, Christensen, and Diewert 1982; Malmquist 1953). This nonparametric method satisfies the challenging requirements of measuring MCP because it captures the multiple objectives of marketing communications in multiple periods and without a prior assumption about the response function. Unlike most regression approaches that offer a comparison with average performers under the assumption of central tendencies, the Malmquist technique benchmarks each firm's productivity against the best practices that perform above average (Färe et al. 1992). In addition, its dynamic time-series nature accommodates the lagged, carryover effects of advertising and sales promotion. Our application of the Malmquist time-series technique not only extends prior work on cross-sectional and time-invariant counterpart models (cf. Horsky and Nelson 1996; Murthi, Srinivasan, and Kalyanaram 1996) but also provides a direct answer to a series of calls for scientific measures of marketing productivity (McLaughlim 1997; Rust et al. 2004).

Furthermore, we develop and test a theoretical framework that predicts that firms that are more efficient in converting advertising and promotion investments into sales and corporate reputation may create greater shareholder value over time. Because MCP facilitates the level and speed of future cash flows while reducing their volatility and vulnerability (Srivastava, Shervani, and Fahey 1998, 1999), we suggest that it also should promote shareholder value in terms of the forward-looking market value measured by Tobin's q and stock return.

Another important insight of our framework is that we do not predict a monotonic linear relationship between MCP and firm market value but rather nonmonotonic curvilinear influences of MCP. Because of the possible conflict between productivity and customer satisfaction (Anderson, Fornell, and Rust 1997), the unrestricted pursuit of improvements in MCP may actually harm firm shareholder value. There is a limit to the incremental performance contributions of MCP, and departing from the optimal point (too low or too high increases of MCP) may lead to reduced Tobin's q and stock return. We also hypothesize about why and how the impact of MCP changes across firms with different levels of research and development (R&D) investment and competition intensity.

Next, we present the hypotheses development for the impact of MCP and the moderated relationships, followed by details of our suggested process for measuring MCP. Using secondary longitudinal data from *Fortune*'s large companies, we find that MCP is positively related to Tobin's q and stock return. For practitioners, this finding helps justify advertising and sales promotion budgets and supports marketing's long-term credibility. However, the supported inverted U-shaped relationship between MCP and firm market value suggests some "dangers" of blindly improving MCP. Our work represents a first step in measuring and valuing marketing communication investments, which must be taken before the productivity crisis can be reversed.

MCP and Shareholder Value

Neoclassical production theory in economics posits that, in general, productivity represents the conversion of economic inputs (e.g., labor, capital) into desirable outputs (i.e., sales, profits) or the ratio of the outputs of an economic activity to the inputs required by that activity. According to Caves, Christensen, and Diewert (1982) and Farrell (1957), in a production setting, managers attempt to improve the value of a firm by operating it in the most productive manner. That is, managers maximize the firm's desirable outputs or outcomes with the minimal amount of inputs or resources. This definition of productivity in an economic production context has been well accepted in marketing (Bucklin 1978; Murthi, Srinivasan, and Kalyanaram 1996; Sevin 1965).

Following this stream of research, we define MCP as the optimally weighted ratio of marketing outputs (sales level, sales growth, and corporate reputation) to marketing communication expenditure (advertising media spending in broadcast, print, and outdoor and sales promotion expenditure).[2]

---

[1] Using the backward-looking performance measure of ROI, the impact of advertising is mixed, according to the literature (e.g., positive: Chauvin and Hirschey 1993; Mizik and Jacobson 2003; negative: Erickson and Jacobson 1992; Holstius 1990).

Figure 1 presents our theoretical framework of MCP and shareholder value. This framework provides a base from which we predict (1) how MCP adds to a firm's shareholder value in a curvilinear manner and (2) the moderating role of contextual factors (R&D and environmental competition intensity) for returns on MCP.

Curvilinear Influences of MCP

In this section, we offer the logic for the dynamic impact of MCP and predict that MCP has nonlinear influences on shareholder value. That is, an initial increase of MCP will enhance the forward-looking performance of Tobin's q and stock returns, but beyond an optimal point, further increases in MCP will become harmful.

An initial increase of MCP over time is positively associated with a firm's market return (i.e., Srivastava, Shervani, and Fahey 1999). Marketing communication productivity can accelerate cash flows in several ways. Efficient marketing communication investments offer the same levels of sales and brand reputation with less advertising and promotion spending. The saved expenditures, which represent extra working capital, then can be invested in other high net present value projects that address long-term performance (Anderson, Fornell, and Rust 1997; Mittal et al. 2005). Over time, this cost saving and higher productivity may accelerate cash flows (Srivastava, Shervani, and Fahey 1999). Furthermore, because the time for market acceptance may be reduced through learning and the persuasion effects of productive and successful advertising, the product adoption and diffusion rate might be facilitated (Erickson and Jacobson 1992; Vakratsas and Ambler 1999). In addition, superior marketing communication programs build customer equity, which leads to higher future repurchases and margins (Keller 1998; Keller and Aaker 1993; Rust 1986). Indeed, productive advertising and promotion help promote product awareness, initial product trials, and repeated purchases (Chauvin and Hirschey 1993; Luo and Donthu 2001; Szymanski, Bharadwaj, and Varadarajan 1993). From another perspective, investors may regard an improvement in efficient marketing and advertising expenditures as good news in the financial markets, which can have a positive impact on the long-term implications of future cash flows (e.g., Hiemstra and Jones 1994). Furthermore, productive advertising and sales promotion programs may reduce the vulnerability and volatility of cash flows through reduced brand-switching intentions and decreased customer churn rates (Reichheld 1996; Vakratsas and Ambler 1999). Customers who retain a higher brand identity with the firm tend to be more responsive to future advertising and promotions. Dynamically, this tendency may lead to reduced market risks for the firm (Srivastava, Shervani, and Fahey 1999). Because MCP results that are too low and employed with inefficient and unproductive marketing communications investments mean overspending in advertising and promotions (Aaker and Carman 1982; Rust et al. 2004), the wasted and inefficient resources likely hamper the speed and level of cash inflows to the firm.

Thus far, we have argued that increasing MCP positively affects firms' market returns. However, we expect that this impact may not be strictly linear. Just as in the case of too low MCP with overspending, an extremely high increase of MCP with minimum spending in advertising and promotion programs may not be optimal. As Anderson, Fornell, and Rust (1997, p. 129) point out, because firms that try to achieve high productivity (i.e., reduce labor costs, cut employee training expenditures) experience lowered levels of customer satisfaction, their "future profitability may be threatened." This potential conflict between productivity and customer satisfaction suggests that the unrestricted pursuit of high productivity and efficiency through relentlessly "lean" marketing and advertising expenditures can lead to a backlash that involves compromises and hindrances in customer equity efforts, initiatives to achieve premeditated levels of customer awareness and corporate exposure, and new product innovations (Dutta, Narasimhan, and Rajiv 1999; Szymanski, Bharadwaj, and Varadarajan 1993). These consequences eventually lead to diminished future performance returns of MCP. In addition, because firms tend to enhance productivity by first cutting the expenses that are most wasteful, subsequent efforts to improve MCP by further cutting advertising and promotion expenditures may cause harm over time and lead to reduced market returns. In other words, there is a limit to the positive incremental return to MCP. At the margin, increasing MCP beyond an optimal point may eventually result in negative returns. This line of reasoning for the curvilinear influence of MCP is also supported by recent empirical accounting and finance research, which finds nonlinear relationships between market value and various predictive variables (Byrd and Hickman 1992; Hiemstra and Jones 1994; Ng 2005; Sougiannis 1994). For example, Ng (2005) finds a curvilinear relationship between corporate ownership and Tobin's q. According to Hiemstra and Jones (1994), trading volume has a nonlinear influence on stock returns. Indeed, Hiemstra and Jones argue (p. 1642) that "changes in stock market volatility can only be properly modeled with nonlinear models."

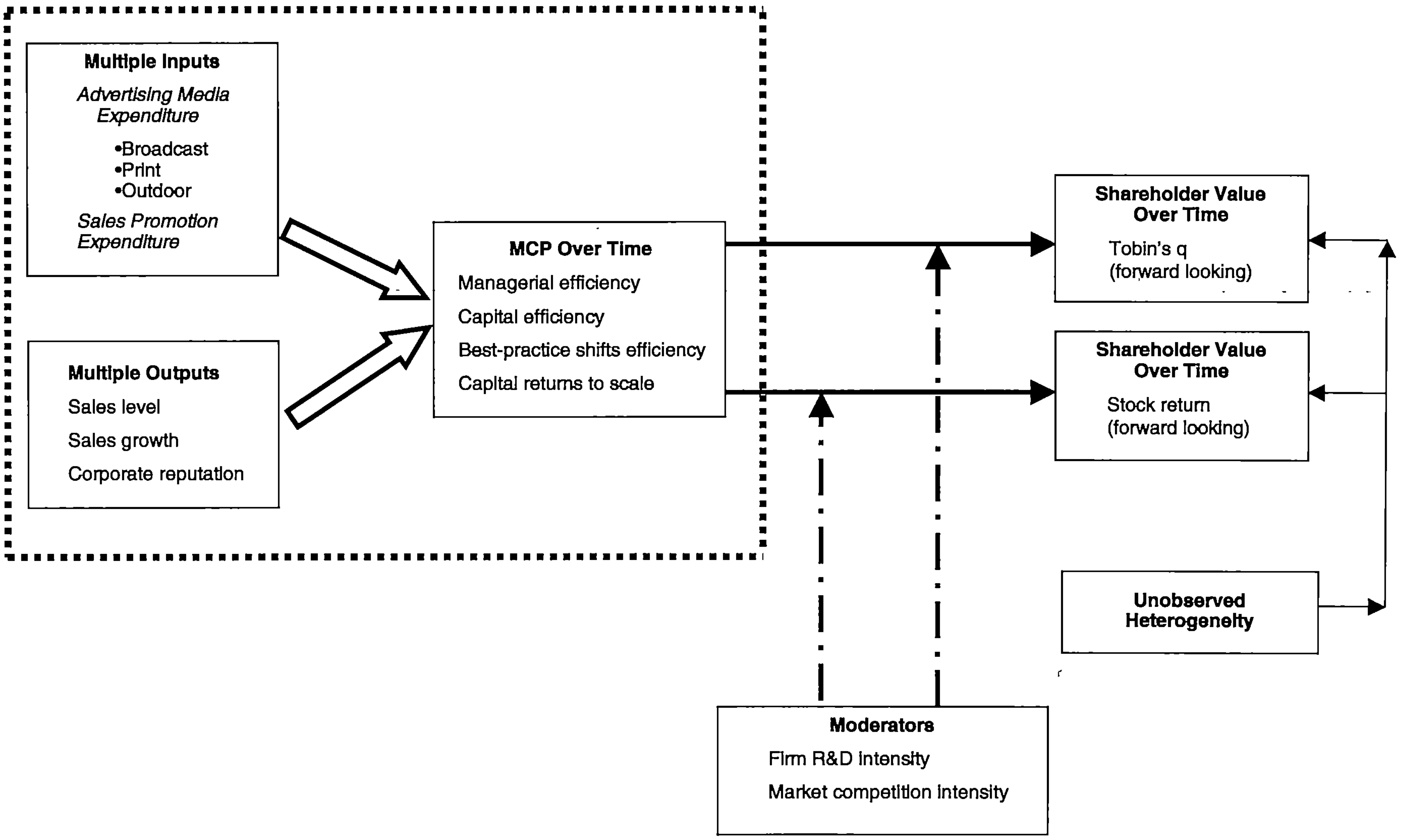
Taken together, this discussion suggests that departing from an optimal point—that is, increases in MCP that are too low or too high—leads to decreases in the forward-looking market value, exhibiting an inverted U-shaped relationship.

> H_1: MCP has an inverted U-shaped influence (positive linear but negative quadratic coefficients) on the forward-looking performance measures of Tobin's q and stock return.

---

[2] Our MCP modeling approach fits with Penrose's (1959) original work on the RBV, which argues that the process of converting resources determines a firm's competitive advantage more than does the level of resources. In addition, because the RBV pertains to best performances and above-average rents (Hansen, Perry, and Reese 2004), our model coincides with the RBV neatly in that it compares each firm with the best-practices frontiers. Our MCP measures examine the productivity (not level) of advertising and promotion expenditure.

FIGURE 1
A Framework of MCP and Shareholder Value



The Moderating Role of R&D Intensity and Competition

We expect that the association between MCP and shareholder value may not be universal but rather may change according to the organizational context. We turn to the RBV perspective (Barney 1986; Penrose 1959; Wernerfelt 1984) and Srivastava, Shervani, and Fahey's (1998) market asset framework to gather the logic surrounding some moderators of this association.

A firm's R&D intensity may change the strength of the association between MCP and market value. In the literature, R&D and marketing are deemed to be complementary resources (Dutta, Narasimhan, and Rajiv 1999; Srivastava, Shervani, and Fahey 1998). Because the RBV considers firms different bundles and configurations of strategic resources, the synergistic interaction between superior marketing and R&D may create additional long-term value, beyond the direct impact of efficient marketing programs. Productive marketing and advertising investments likely obtain valuable feedback from the market and customers, which facilitates R&D and new product innovations. In turn, this offers enhanced cash flow to the firm. In addition, intense R&D in the new product development process can "ensure speedy and successful commercialization of technologies and products at a low cost" (Dutta, Narasimhan, and Rajiv 1999, p. 551; see also Gupta, Raj, and Wilemon 1987). Thus, superior MCP coupled with higher R&D intensity over time may lead to even greater speed and levels of cash flow, along with lower vulnerability and volatility, which can promote greater market value in the long run (Griffin and Hauser 1996; Srivastava, Shervani, and Fahey 1999). Working from the RBV perspective, Dutta, Narasimhan, and Rajiv (1999) find that R&D capability facilitates the positive influence of marketing capability and brings complementary rents into the firm dynamically, beyond the contribution of superior marketing efficiency alone.

In summary, on the basis of the RBV and prior findings, we suggest that in conjunction with higher R&D intensity in the firm over time, MCP leads to greater firm market value.

> H_2: MCP has a stronger influence on the forward-looking performance measures of Tobin's q and stock return in firms with higher (versus lower) R&D intensity.

Finally, we predict that environmental competition intensity expands the complementary rents of MCP and intense R&D, as we state in H_2. In particular, in a competitive market, buyers may choose from more substitutes of products and services. To lock in buyers, firms strategically build brand loyalty and create exit barriers that facilitate the speed and level of their cash flows. Similar to Schumpeter's (1934) notion of continuous entrepreneurial innovations due to market competition, in more competitive markets, firms are motivated (reactively or proactively) to leverage not only efficient advertising programs but also R&D investments over time to dynamically upgrade and renew cash-cow brands to shield themselves from head-to-head competition for more entrepreneurial rents. The more competitive the market, the more firms may rely on marketing and R&D skills simultaneously to ensure the successful commercialization of their existing and new technologies, products, and services (i.e., Griffin and Hauser 1996; Mizik and Jacobson 2003).

In short, the complementary effect of marketing and R&D (Dutta, Narasimhan, and Rajiv 1999) may be expanded in more competitive markets, which would lead to greater cash flows and firm market value. Therefore, it is likely that MCP has the strongest influence on Tobin's q and stock returns when firms have higher R&D intensity *and* operate in highly competitive markets.[3]

> H_3: MCP has the strongest influence on the forward-looking performance measures of Tobin's q and stock return in firms that have high R&D intensity *and* are in more competitive markets.

Measuring MCP: The Modeling Process

We now detail a process of modeling MCP that is based on the time-series Malmquist approach. We present a flow chart of this process in Figure 2. As we show in the flow chart, there are two significant stages involved in our modeling process. Stage 1 focuses on measuring MCP, and Stage 2 focuses on valuing MCP (as we detail in the "Analyses and Results" section). In Stage 1, we first define MCP as the conversion ratio of marketing communication inputs to outputs, which requires a logical identification of the multiple inputs to multiple outputs.

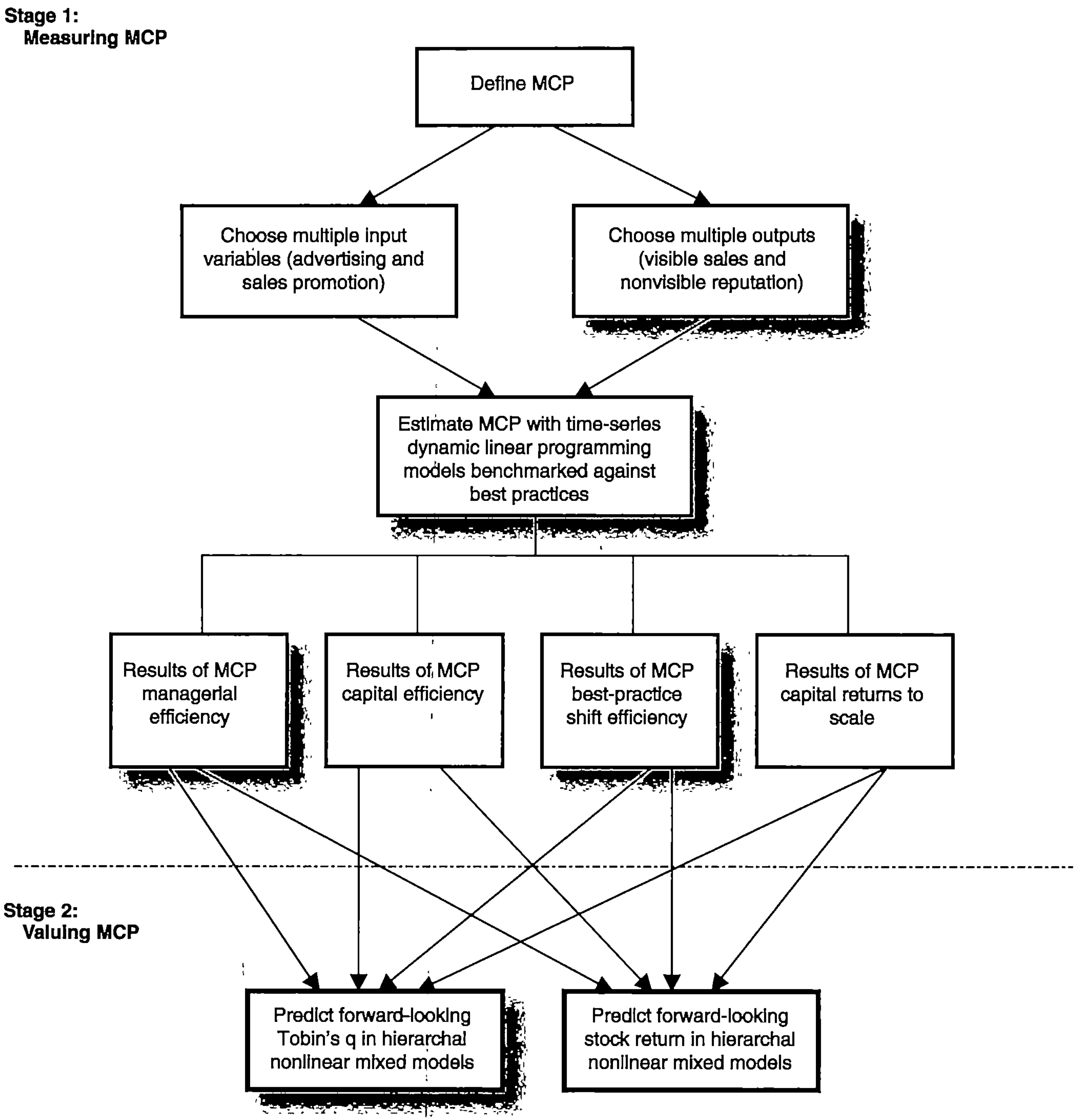
For the inputs of MCP, we focus on two main categories: advertising media spending and sales promotion expenditures. Advertising media spending includes three dimensions: expenditures in broadcast, print, and outdoor. Specifically, a firm's advertising expenditures in broadcast media equals the sum of television and radio expenditures, including network television spots on ABC, CBS, NBC, Fox, WB, UPN, and Pax TV and television spots on 575-plus stations in the top 100 U.S. markets, which encompasses more than 180 programs per month and 37 cable networks, as well as four radio networks and spot radio spending on 4000 stations in more than 225 markets. The MCP input of advertising expenditures in print media equals the sum of newspaper, magazine, and other expenditures, including more than 140 newspapers in 60 of the top U.S. markets, more than 230 consumer magazines, and Sunday magazines. The MCP input of advertising expenditures in outdoor media is the total expenditures in more than 300 outdoor plant operator markets. Finally, the MCP input of sales promotion expenditures is the sum of spending in sales promotions, direct mail, co-op spending, coupons, catalogs, business and farm publications, and special events.

Next, we specify three outputs of MCP—sales level, sales growth, and corporate reputation—because marketing communication expenditures influence these outcomes simultaneously. Note that measuring the sales level requires an approach to control the influence of other relevant variables and carryover effects of advertising and marketing

---

[3] Although the main effects of competition may be positive or negative, we investigate only the moderating effect.

FIGURE 2
MCP Methodology Flow Chart


Stage 1:
Measuring MCP
Define MCP
Choose multiple input variables (advertising and sales promotion)
Choose multiple outputs (visible sales and nonvisible reputation)
Estimate MCP with time-series dynamic linear programming models benchmarked against best practices
Results of MCP managerial efficiency
Results of MCP capital efficiency
Results of MCP best-practice shift efficiency
Results of MCP capital returns to scale
Stage 2:
Valuing MCP
Predict forward-looking Tobin's q in hierarchal nonlinear mixed models
Predict forward-looking stock return in hierarchal nonlinear mixed models

expenditures (i.e., Kim, Bhargava, and Ramaswami 2001; Vakratsas and Ambler 1999), as we discuss subsequently. The MCP output of sales growth is the three-year compounded annual growth rate of sales in the primary Standard Industrial Classification of the firm (Rao, Agarwal, and Dahlhoff 2004). The MCP output of corporate reputation is defined as the firm's overall appeal to the public and is based on its historical actions and future prospects (McGuire, Schneewels, and Branch 1990; Roberts and Dowling 2002). According to *Fortune* (2005), this variable assesses the perceived reputation of the large U.S. firms and is based on the research of more than 10,000 financial analysts and investors.

After we specify the inputs and outputs of MCP, we introduce the dynamic Malmquist approach to estimate MCP. The intuition of this approach is that firms attempt to consume the least possible amount of inputs to achieve the same level of desired outputs from time t to time t + 1. If a firm cannot reduce its inputs (i.e., use fewer resources) without hurting its output levels, it is considered productive over time. Otherwise, it is unproductive and inefficient. Thus, productive firms as a whole represent the so-called best-practice frontier that dynamically benchmarks each firm's MCP in the transformation of inputs into outputs (Färe et al. 1992; Luo 2004).

To calculate MCP from t to t + 1 for each firm, Malmquist (1953) initially developed dynamic models to assess the total factor productivity of general economic activities over time. Later, Färe and colleagues (Färe et al. 1992; Färe, Grosskopf, and Lovell 1994) constructed the time-series linear programming (or data envelopment analysis [DEA] –based) Malmquist productivity index.[4] As we show in the Appendix, this measure of MCP for each firm can be calculated using two single-period and two mixed-period efficiencies based on DEA programming (Banker, Charnes, and Cooper 1984; Charnes, Cooper, and Rhodes 1978). To the best of our knowledge, the literature has provided no time-series applications of the DEA-based Malmquist approach to measure and value the various components of MCP (for technical details and numeric examples of DEA, see Horsky and Nelson 1996; Luo and Donthu 2005; Murthi, Srinivasan, and Kalyanaram 1996; see also the Appendix).

The Malmquist approach is well suited for measuring MCP and tracking its changes over time because of its methodological advantages. First, it can estimate the productivity of marketing communication expenditures "without a priori information on tradeoffs among inputs and outputs" (Chen and Ali 2004, p. 239). This is advantageous because the function of market responses to advertising and promotions is conflicting and inconclusive (Strong 1925; Vakratsas and Ambler 1999). In addition, some advertising expenditures are intended to promote corporate brands, whereas others are intended to increase sales revenue, and still others aim for a little bit of both. Therefore, without prior knowledge about which part of the advertising and marketing expenditure produces which part of the outputs, the Malmquist approach is very suitable. Second, unlike regression techniques that use one output at a time (Caves, Christensen, and Diewert 1982), Malmquist models are not limited by the number of outputs, which is desirable for measuring MCP because marketing communications can achieve multiple outcomes (i.e., visible outcomes, such as sales, and invisible outcomes, such as reputation) at the same time.[5] Third, the time-series nature of Malmquist modeling offers an advantage for benchmarking MCP because advertising and sales promotion expenditures have lagged and carryover effects; outcomes may unfold over a long period in the marketplace (Färe et al. 1992).

The Malmquist approach also fits well with the RBV of the firm (e.g., Barney 1986; Penrose 1959). According to Penrose (1959), a firm's best practices in the process of converting resource inputs to outputs (or MCP in our context) ultimately determine firm performance, not just the levels of those inputs. Hansen, Perry, and Reese (2004, p. 1280) recently acknowledged that "RBV is fundamentally a theory about extraordinary performers or outliers. The statistical methods used in applying the theory must account for individual firm differences, but should not be based on means, which statistically neutralize firm difference." The Malmquist approach is just such a method and identifies best-practice frontiers and benchmarks MCP at the individual level.

The final step in Stage 1 is to decompose the overall MCP into several subcomponents across firm-year observations (see the Appendix). Intuitively, the first dimension of MCP reflects the change of MCP from time t to t + 1 in terms of managerial efficiency, which we label MCP1 (i.e., the effects of moving toward the best-practice frontiers). This efficiency represents how much improvement in MCP can be achieved if managers enhance their skills of managing marketing communication programs compared with the best practices. Another dimension is the change of MCP from time t to t + 1 in terms of capital efficiency, which we label MCP2 (i.e., the effects of moving between the best-practice frontiers with different returns-to-scale model specifications). This efficiency represents how much improvement in MCP can be achieved if managers can promote the economies of scale when using marketing communication capitals. It also offers a dimension that represents the change of MCP from time t to t + 1 in terms of best-practice shift efficiency, which we label MCP3 (i.e., the effects of moving away from the best-practice frontiers from t to t + 1). This efficiency indicates how much improvement in MCP can be achieved through factors that are unrelated to marketing communication managerial skills but are associated with intertemporal, uncontrollable variables, such as industry changes, communication innovations, and environmental variables. Finally, the dimension of returns to scale of marketing communications capital (i.e., the effects of moving along the best-practice frontiers) is a categorical variable that can be increasing, constant, or decreasing. We create two dummy variables (MCPd1 and MCPd2) for these three kinds of returns to scale.

Next, we describe the multiple sources of secondary data we collected to measure MCP. Then, we present the analysis approach in Stage 2, which investigates the value of MCP by relating its subdimensions to Tobin's q and stock returns.

Data and Measures

We obtained data on the advertising and promotions expenditures, sales revenue, corporate reputation, Tobin's q, stock return, accounting profitability, and firm characteristics variables from multiple archived sources. Specifically, we collected secondary data from four sources: COMPUSTAT, Competitive Media Reporting (CMR), the *Fortune* reputation index, and the Center for Research in Security Prices (CRSP). To evaluate the contribution of MCP to firms' market value, we selected the large publicly traded *Fortune* 1000 companies (and the top marketers and advertisers) for eight consecutive years (1995–2002). We report the details of the cross-sectional time-series databases and the variables in Table 1.

---

[4] Compared with parametric alternatives such as stochastic frontier, thick frontier, distribution-free analyses, and translog specification, the Malmquist linear programming estimates are more stable in Monte Carlo simulation studies (e.g., McAllister and McMannus 1993) and offer richer managerial implications (e.g., Luo 2004; Luo and Donthu 2005; Wheelock and Wilson 1999).

[5] Although multivariate analysis methods such as canonical correlation can deal with multiple outputs and inputs, they are essentially parametric approaches that assume normal distribution and central tendency (i.e., they compare with average performers rather than best practices). In addition, canonical correlation can estimate only one set of coefficients across all firms and thus produce, at best, only one index of MCP for each firm (cf. the Malmquist approach offers a more complete understanding of different facets of MCP). We thank an anonymous reviewer for highlighting this point.

To estimate MCP over time, we need data for the four marketing communication expenditure inputs (i.e., in broadcast, print, outdoor, and sales promotion) and three outputs (sales level, sales growth, and corporate reputation). The MCP input variable data come from the CMR archives, which offer fairly comprehensive information about the breakdown of advertising and promotion expenditures by the large firms in top markets (e.g., Anderson, Fornell, and Mazvancheryl 2004; Rao, Agarwal, and Dahlhoff 2004). Although COMPUSTAT also offers information on advertising and sales expenditures, many data points are missing because firms are not required to report this type of expenditure (see Joseph and Richardson 2002). In addition, unlike the CMR data set, COMPUSTAT does not break down advertising expenditure data into different types of media spending.

For MCP output variables, we first used sales level that has controlled the carryover effects of advertising (i.e., Chauvin and Hirschey 1993; Vakratsas and Ambler 1999) and the impact of other relevant factors, such as R&D, firm size, and competition.[6] Specifically, we employed the following functions:

$$(1) \qquad \text{SALES}_t^{\sim} = \text{SALES}_t - \text{SALES}_{t\_\text{predicted}}, \text{ where}$$

$$\begin{aligned}\text{SALES}_{t\_\text{predicted}} = {} & a + b_0\text{SALES}_{t-1} + b_1 \times \text{iADV}_{t-1} \\ & + b_2 \times \text{RD}_t + b_3 \times \text{RD}_{t-1} + b_4 \times \text{SIZE}_t \\ & + b_5 \times \text{SIZE}_{t-1} \\ & + b_6 \times \text{Focus of FIRM}_t \\ & + b_7 \times \text{Focus of FIRM}_{t-1} \\ & + b_8 \times \text{competition}_t + b_9 \times \text{competition}_{t-1} \\ & + e.\end{aligned}$$

COMPUSTAT provided data for the book value of total assets, sales revenue, R&D intensity, firm size, the focus of the firm, industry competition intensity, and other such variables. Firm size is the log of the number of employees, and focus of the firm is the number of industry segments in which the firm operates (Rao, Agarwal, and Dahlhoff 2004). In line with prior literature in marketing and management (i.e., Anderson, Fornell, and Mazvancheryl 2004; Fischer and Pollock 2004), we measured industry competition intensity using the Herfindahl concentration index, which we derived from COMPUSTAT. A low degree of concentration indicates that the marketplace is fragmented and competition intensity is high. We measure R&D intensity as the ratio of R&D spending to total assets for each firm-year observation (Anderson, Fornell, and Mazvancheryl 2004), which we also derived from the COMPUSTAT database.

The second output variable is sales growth, which we also derived from the COMPUSTAT data set and calculated as the three-year compounded annual sales growth rate rather than a simple growth rate, such as sales_t versus sales_{t-1} (Rao, Agarwal, and Dahlhoff 2004). Sales growth could be measured with a simple two-year-based growth rate (t versus t – 1). However, to be more parsimonious, we measured sales growth on the basis of the three-year compounded rate (t versus t – 1 and t – 2). Because this compounded rate is essentially a "smoothed" rate of growth, it has relatively fewer concerns, such as the possible existence of some random dramatic shocks. In general, the compounded annual sales growth rate (t versus t – 1 and t – 2) is preferred to a simple growth rate (t versus t – 1) in both marketing and accounting literature (see Rao, Agarwal, and Dahlhoff 2004).

We obtained data for the final MCP output variable, corporate reputation, from the annual reputation survey of the most admired U.S. corporations (Houston and Johnson 2000; Roberts and Dowling 2002). This measure of perceived overall reputation of the largest U.S. firms uses an 11-point scale (0–10) to assess eight indicators (innovativeness, employee talent, social responsibility, long-term investment value, financial soundness, use of corporate assets, quality of management, and quality of products and services; see *Fortune* 2005). *Fortune* surveys more than 10,000 executives, directors, and securities analysts every year.

Many prior studies have successfully used this corporate reputation score in both marketing and strategy (Houston and Johnson 2000; Roberts and Dowling 2002). In particular, McGuire, Scheeweis, and Branch (1990, p. 170) note that the "*Fortune* reputation is one of the *most comprehensive* and widely circulated surveys of attributes available [for measuring reputation]. Both the quality and number of respondents are comparable or superior to the 'expert panels' usually gathered for such purposes." Thus, research across disciplines seems to accept this best-available measure of corporate reputation. However, because it has been noted that the *Fortune* reputation measure depends on financial performance (e.g., Aaker and Jacobson 2001), we also address this reverse-causality issue in our analyses, as we discuss subsequently, using the approach that Roberts and Dowling (2002) recommend.[7]

---

[6] Conceptually, it is important to have a "clean" measure of sales level when examining the relevant contribution of marketing expenses at a given point of time. That is, we need to tease out the impact of many other variables and carryover effects of marketing communication expenditures. Dutta, Narasimhan, and Rajiv (1999) use a similar approach to adjust their output variable when evaluating capability efficiency (MCP productivity in our context).

[7] To control the reverse-causality issue between financial performance and corporate reputation (Aaker and Jacobson 2001), we adopted the approach that Roberts and Dowling (2002) recommend. As we discussed in the "Additional Analyses" subsection, our results hold after accommodating this reverse-causality concern. We thank an anonymous reviewer for this point.

TABLE 1
Measures and Data Sources

Measure	Operationalization	Data Sources	Period
MCP	A longitudinal linear-programming-based ratio of multiple outputs (sales revenue, market growth, and corporate reputation) to multiple inputs (broadcast, print, outdoor, and sales promotion expenditures)		1995–2002
Forward-looking performance of Tobin's q	The result of ([share price × number of common stock outstanding + liquidating value of the firm's preferred stock + short-term liabilities – short-term assets + book value of long-term debt]/book value of total assets)	Derived from CRSP and COMPUSTAT	1994–2003
Forward-looking performance of stock return	The result of (current year's share price × number of common stock outstanding + dividends – previous year's share price × number of common stock outstanding/[previous year's share price × number of common stock outstanding])	Derived from CRSP and COMPUSTAT	1994–2003
Backward-looking performance of return on assets	The ratio of net income before extraordinary items to assets	Derived from COMPUSTAT	1995–2002
Sales level	Adjusted for advertising and marketing carryover effect, R&D contemporary and carryover effect, competition justification effect, and firm-size effect	Derived from COMPUSTAT	1994–2003
Sales growth	The compounded annual sales growth rate over the previous three years	Derived from COMPUSTAT	1993–2003
Corporate reputation	*Fortune* reputation index (corporate reputation based on an annual survey of the most admired U.S. corporations)	*Fortune*	1995–2002
Advertising broadcast expenditure	The sum of television and radio expenditures, such as in network television spots on ABC, CBS, NBC, Fox, WB, UPN, and Pax TV and spot television on 575-plus stations in the top 100 U.S. markets; more than 180 syndicated television programs per month, 37 cable networks, and four radio networks and spot radio spending from 4000 stations in more than 225 markets	CMR	1995–2002
Advertising print expenditure	The sum of newspaper, magazine, and print other expenditures including 140-plus newspapers in 60 of the United States' top markets; monitors space in national newspapers, including *The Wall Street Journal*, *USA Today*, and *The New York Times*; 230-plus consumer magazines and Sunday magazines	CMR	1995–2002
Advertising outdoor expenditure	The sum of more than 300 outdoor plant operator markets	CMR	1995–2002
Sales promotion expenditure	The sum of spending in sales promotion, direct mail, co-op spending, couponing, catalogs, business and farm publications, and special events	CMR	1995–2002
Competition intensity	Herfindahl concentration index (derived by using the lagged sales for all the companies in the same two-digit Standard Industrial Classification code for each firm-year observation)	Derived from COMPUSTAT	1994–2002
R&D intensity	R&D expenditure/book value of total asset_{t-1}	COMPUSTAT	1994–2002
Firm size	Number of employees	COMPUSTAT	1995–2002
Focus of the firm	Number of industry segments the firm operates	COMPUSTAT	1995–2002

Because we tried to merge data from these different sources and each source has some missing values, our sample size was reduced. After searching other sources, such as Standard & Poor's industry reports, *AdWeek*, Moody's report, Compact Disclosure, and company annual reports, we obtained complete data on 89 companies. As a result, we have 712 data points across the firms from 1995 to 2002.

In line with the work of Jacobson and colleagues (Aaker and Jacobson 1994, 2001; Mizik and Jacobson 2003), we derived the forward-looking performance measures of stock return using both the COMPUSTAT and the CSRP databases. Tobin's q, which is comparable across industries (Anderson, Fornell, and Mazvancheryl 2004; Dutta, Narasimhan, and Rajiv 1999), offers another forward-looking measure of firm market value. We followed prior marketing studies (Lee and Grewal 2004; Rao, Agarwal, and Dahlhoff 2004) to calculate Tobin's q for each firm-year observation. In computing the stock return and Tobin's q, we also controlled for the influence of return on assets (ROA; see Chauvin and Hirschey 1993; Erickson and Jacobson 1992), which we measure as the ratio of net income before extraordinary items to assets, which we derived from COMPUSTAT.

Analyses and Results

Results of Measuring MCP

Because we wanted to benchmark and track the longitudinal changes of several different dimensions of MCP over time, we estimated 39,872 linear programs (4 efficiency models × 2 returns to scale × 712 data points × [8 – 1] years) to derive the MCP components for the sample. In addition, following Färe, Grosskopf, and Lovell's (1994) approach, we ran 22,784 additional linear programming models to determine the returns to scale of marketing expenditure.[8] We summarize the results of the subdimensions and the overall MCP in Table 2. The interpretation of the results is that when MCP > 1 (for MCP dimensions and overall MCP), MCP improves from t to t + 1; when MCP = 1, no change occurs; and when MCP < 1, MCP decreases from t to t + 1.

Our data suggest that MCP improved over time, as we show in Table 2. Overall MCP was stable during 1995–1998, but it increased steeply during 1998–2001. For the individual dimensions of MCP, we observed an overall trend of improved managerial efficiency (MCP1) and dramatic increases (with a peak) in 2000–2001. The capital efficiency of marketing communications expenditure (MCP2) slightly increased during 1995–1999 but then decreased during 1999–2002. In addition, the best-practice shift efficiency (MCP3) seemed to increase over time but decreased during 1997–1998 and 1999–2000. Therefore, both managerial and best-practice shift efficiency seem to be the drivers of the improvement of the firms' overall MCP over time in this sample.

We also break down the overall results of MCP according to the different contextual factors in terms of firm differences in R&D spending and market competition intensity. As we present in Figure 3, Panel A, firms with high R&D intensity tend to have sustained, superior MCP during the 1995–2002 period, especially between 2000 and 2001. There is a significant difference between the high- and the low-R&D groups (with a median split) in their overall MCP changes over time (t = 4.82, $p < .05$). In support of the literature that suggests that high R&D intensity is important for firm performance (e.g., Dutta, Narasimhan, Rajiv 1999; Gupta, Raj, and Wilemon 1987; Rust, Moorman, and Dickson 2002), this finding indicates that it is also positively related to an intermediary marketing outcome of MCP. Furthermore, it supports the recent theoretical argument about the need to investigate intermediary measures of marketing productivity (Rust et al. 2004).

For market competition intensity, as we show in Figure 3, Panel B, the results indicate a significant difference between the high- and the low-competition groups (with a median split) in their overall MCP changes over time (t = 5.29, $p < .05$). Although low competition intensity is associated with higher overall MCP for the most part, during 2000–2002, high market competition is related to improved overall MCP. Because Malmquist MCP results are linear-programming-based nonparametric results, we report both Spearman and Pearson correlation results of the variables in Table 3.

Results of Valuing MCP

Hypotheses testing approach. To test H_1–H_3, we employ a hierarchal nonlinear mixed model to control for unobserved heterogeneity and unmeasured nuances in the effects of MCP on shareholder value. Prior studies suggest that for cross-sectional time-series data, large variance may be due to the unobserved heterogeneity in firm-level factors, such as marketing managerial expertise, and time-level factors, such as changes in consumer learning and industry trends over time (Boulding and Staelin 1993; Jacobson 1990). In particular, Jacobson (1990, p. 94) warns that "models and methods that ignore unobservable factors and disregard the consequences of neglecting their effects … generate insidious results." However, because there is no single ideal model recommended to control for the unobservable effects (e.g., Baltagi 2001; Boulding 1990; Grilliches and Hausman 1986; Murthi, Srinivasan, and Kalyanaram 1996), we specify the following nonlinear mixed model to account for two-way (firm and year) random effects with the dependent variables of Tobin's q:[9]

$$
\begin{aligned}
(2)\qquad \text{Tobin's } q_{it} = {} & \mu + \beta_1 \times \text{MCP1}_{it} + \beta_2 \times \text{MCP2}_{it} \\
& + \beta_3 \times \text{MCP3}_{it} + \beta_4 \times \text{MCPd1}_{it} \\
& + \beta_5 \times \text{MCPd2}_{it} + \beta_6 \times \text{CPI}_{it} + \beta_7 \times \text{RD}_{it} \\
& + \beta_8 \times \text{ROA}_{it} + \Sigma\beta_{\text{inter}}\text{Interactions}_{it} \\
& + \Sigma\beta_{\text{squar}} \times \text{MCP-squared}_{it} \\
& + \Sigma\beta_{\text{intsq}} \times \text{MCP-squared} \times \text{Interactions}_{it} \\
& + \Sigma\beta_{\text{cova}}\text{Covariates}_{it} + \varepsilon_{it} + \zeta_i + \Omega_t,
\end{aligned}
$$

---

[8] We need only two dummy variables for the three different returns to marketing communications expenditures: increasing, constant, and decreasing. The decreasing-returns case is the base in the dummies.

[9] We also tried to specify the model with cubic terms of MCP dimensions but failed to find significant coefficients.

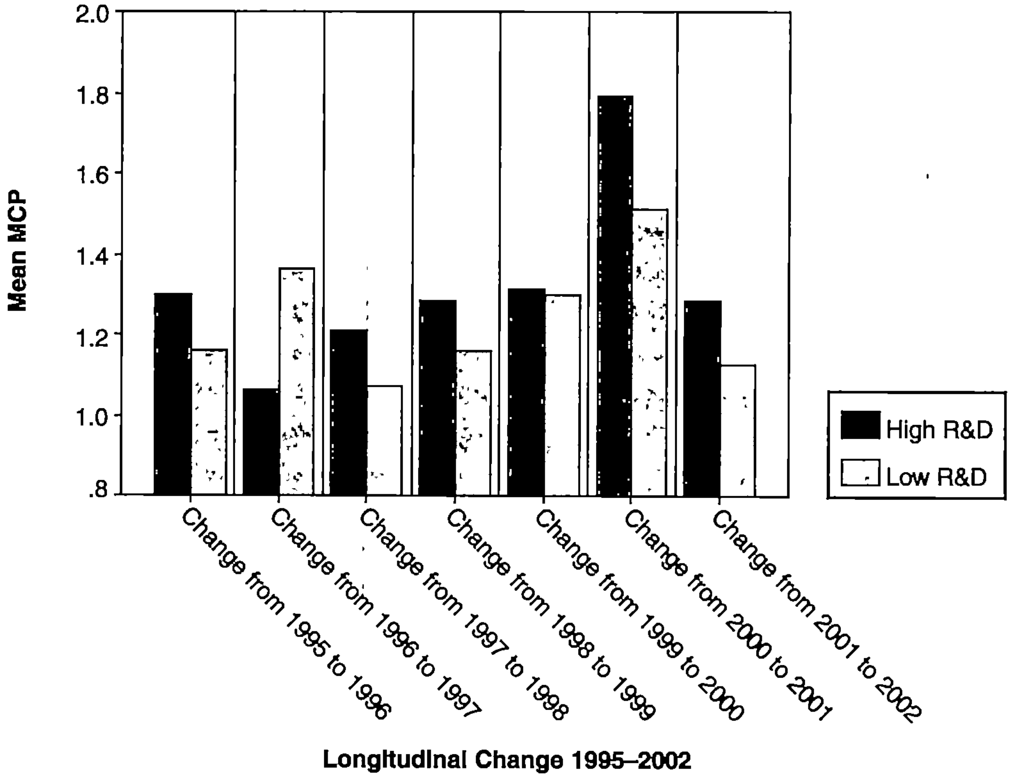
TABLE 2
MCP Summary Statistics

	Overall Productivity Change: 1995–2002	Productivity Change: 1995–1996	Productivity Change: 1996–1997	Productivity Change: 1997–1998	Productivity Change: 1998–1999	Productivity Change: 1999–2000	Productivity Change: 2000–2001	Productivity Change: 2001–2002
MCP: Overall	1.282 (.823)	1.231 (.699)	1.213 (.754)	1.141 (.603)	1.223 (.604)	1.307 (.642)	1.652 (1.479)	1.207 (.480)
MCP1: Managerial efficiency	1.106 (.599)	1.204 (.515)	1.041 (.560)	1.212 (.760)	1.093 (.288)	.948 (.324)	1.301 (.996)	.939 (.290)
MCP2: Capital efficiency	1.012 (.244)	1.014 (.190)	1.022 (.288)	1.018 (.213)	1.070 (.300)	.990 (.117)	.990 (.225)	.983 (.314)
MCP3: Best-practice shift	1.197 (.406)	1.035 (.367)	1.205 (.333)	.961 (.237)	1.094 (.491)	1.386 (.348)	1.283 (.333)	1.412 (.455)

Notes: MCP > 1 indicates that MCP increased from t to t + 1; MCP < 1 indicates that MCP decreased from t to t + 1; MCP = 1 indicates that MCP did not change from t to t + 1. Entries in parentheses are standard errors

FIGURE 3
MCP Longitudinal Changes

A: R&D Breakdown



B: Competition Breakdown

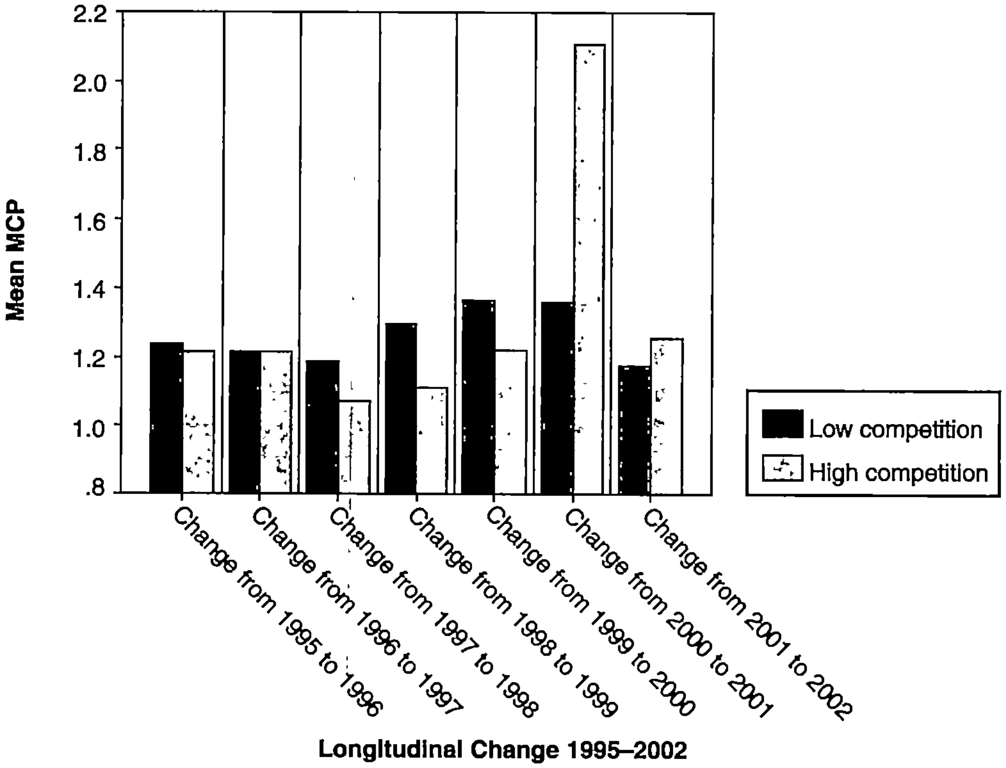


TABLE 3
Parametric and Nonparametric Correlations

	X1	X2	X3	X4	X5	X6	X7	X8	X9
MCP best-practice shift (X1)	1.00	.09*	–.02	.15**	.17**	.03	.18**	.03	.09*
MCP managerial efficiency (X2)	.23**	1.00	.10*	–.12**	.10*	.13**	.09*	.14**	–.01
MCP capital efficiency (X3)	–.09*	.02	1.00	–.04	.14**	.07	.12**	–.03	–.05
MCP returns to scale (X4)	.19**	–.14**	–.02	1.00	.11*	–.04	.10*	.17**	.13**
Tobin's q (X5)	.18**	.10*	.13**	.12**	1.00	.06	.32**	.11**	.11**
ROA (X6)	.06	.11*	.06	–.03	.04	1.00	.18**	.03	.09*
Stock return (X7)	.18**	.10*	.11*	.13**	.35**	.19**	1.00	.10*	.08
Market competition intensity (X8)	.05	.08*	–.06	.15**	.09*	.02	.14**	1.00	.20**
R&D intensity (X9)	.10*	–.07	–.04	.13**	.10*	.10*	.09*	.21**	1.00

*$p < .05$.
**$p < .01$.
Notes: The entries above the diagonal are Pearson parametric correlations, and the entries below the diagonal are Spearman nonparametric correlations.

where

MCP1 = the managerial efficiency dimension of MCP;
MCP2 = the capital efficiency dimension of MCP;
MCP3 = the best-practice shift efficiency dimension of MCP;
MCPd1 = MCP scale dummy 1 with increasing returns to scale;
MCPd2 = MCP scale dummy 2 with constant returns to scale;
MCP-squared = the quadratic terms of MCP1, MCP2, and MCP3;
CPI = competition intensity;
RD = R&D intensity;
ROA = accounting backward-looking performance of ROA;
Interactions_{it} = interaction terms for firm i at time t, or interactions among R&D intensity, competition intensity, and dimensions of MCP; they include all two-way and three-way interactions;
MCP-squared × Interactions_{it} = interaction terms among quadratic terms of MCP1, MCP2, and MCP3, all above two- and three-way interactions for firm i at time t;
Covariates_{it} = covariates for firm i at time t, such as ROA, firm size, focus of the firm, business type, and market service type;
μ = the overall grand intercept;
β_i = the influence of ith explanatory variables;
ε_{it} = the disturbance term, normally and independently distributed with a constant variance σ^2;
ζ_i = the random disturbance of ith firm, constant across periods and normally and independently distributed with a constant unknown variance σ^2_ζ; and
Ω_t = the random disturbance of tth year, constant across firms and normally and independently distributed with a constant unknown variance σ^2_Ω.

We summarize the results in Table 4. The mixed model estimates indicate that for the forward-looking performance of Tobin's q, firm random effects have a substantial intraclass correlation coefficient of 19.68% (8.30/[28.25 + 8.30 + 5.62]). Thus, unobserved/omitted firm factors may account for as much as 19.68% of the variance of Tobin's q, according to the variance structure distribution of the three types of errors in Table 4 (Baltagi 2001). Unobserved time random effects are also significant and account for approximately 13.32% of the variance of Tobin's q. These results imply that to reduce omitted variable bias, an analysis of the cross-sectional time-series panel data should not ignore the issue of unobserved heterogeneity and the influence of omitted variables (e.g., Arellano 2003; Boulding and Staelin 1993).

Hypotheses testing results. The findings provide some support for H_1: MCP has an inverted U-shaped influence on the forward-looking performance. As we show in Table 4, marketing communication managerial efficiency positively influences forward-looking Tobin's q (MCP1: t = 3.87, $p < .01$), and too much improvement in MCP1 negatively influences Tobin's q. The squared term of MCP1 is negative and significant (MCP1-squared: t = –3.51, $p < .01$), thus suggesting an inverted U-shaped relationship. Similarly, marketing communication best-practice shift efficiency has a positive influence on Tobin's q (MCP3: t = 2.86), but too great an increase of this efficiency has a negative influence (MCP3-squared: t = –2.36, $p < .05$), as we expected.

H_2 predicted that MCP would have a stronger positive influence on Tobin's q in firms with higher R&D intensity. As we show in Table 4, when coupled with higher R&D intensity, positive changes in marketing communication managerial efficiency (MCP1 × R&D: t = 2.23, $p < .05$) and best-practice shift efficiency (MCP3 × R&D-related: t = 1.66, $p < .1$) lead to higher Tobin's q for the firm. Thus, H_2 is supported for the Tobin's q model, which not only adds

TABLE 4
MCP and Shareholder Value

Variables	ROA: Backward Looking β	ROA: Backward Looking (t)	Tobin's q: Forward Looking β	Tobin's q: Forward Looking (t)	Stock Return: Forward Looking β	Stock Return: Forward Looking (t)
MCP Dimensions						
MCP1 (managerial efficiency)	7.67	(3.81***)	8.06	(3.87***)	9.28	(4.03***)
MCP2 (capital efficiency)	8.65	(1.86*)	12.21	(3.09***)	12.16	(3.15***)
MCP3 (best-practice shift)	3.73	(1.22)	7.98	(2.86**)	12.73	(3.61***)
MCPd1 (dummy1)	.27	(.98)	5.36	(1.92*)	7.87	(2.52**)
MCPd2 (dummy2)	.13	(.60)	3.55	(1.61)	3.03	(1.31)
Competition intensity (CPI)	8.11	(.83)	13.86	(2.57**)	19.31	(2.82**)
R&D intensity (R&D)	5.62	(1.81*)	7.33	(2.19**)	6.07	(1.85*)
Interactions						
MCP1 × R&D			6.86	(2.23**)	7.22	(2.36**)
MCP2 × R&D	12.45	(3.21***)				
MCP3 × R&D	6.76	(2.62**)	3.32	(1.66*)		
MCP1 × CPI			14.19	(3.57***)	10.42	(2.87**)
MCP2 × CPI			12.11	(2.92**)	13.75	(3.26**)
MCP3 × CPI	4.88	(3.97***)				
MCP1 × R&D × CPI			16.82	(3.84*)	15.20	(3.58***)
MCP3 × R&D × CPI			18.32	(6.18**)		
Curvilinear Terms						
MCP1-squared	16.29	(4.15***)	−15.37	(−3.51***)	−11.76	(−2.21**)
MCP3-squared			−9.28	(−2.36**)		
MCP1-squared × R&D	19.86	(5.86***)	−11.84	(−2.27**)		
MCP1-squared × CPI					−9.45	(−1.83*)
MCP2-squared × CPI			−18.53	(−3.36***)		
MCP1-squared × R&D × CPI			−27.71	(−2.25**)		
ROA	—	—	1.28	(2.31**)	1.87	(4.15***)
Firm size	.01	(.53)	.00	(.22)	.00	(.36)
Markets service types	−.08	(−2.52**)	−.36	(−1.70*)	−.32	(−1.69*)
Business types	−.02	(−1.69*)	−.01	(−.78)	−.03	(−.82)
Focus of the firm	−.03	(−1.23)	−.00	(−.05)	−.04	(−.18)
σ_ε	8.25	(11.32***)	28.25	(35.54***)	26.29	(33.77***)
σ_ζ	1.03	(2.40**)	8.30	(9.33***)	9.12	(11.18***)
σ_Ω	2.16	(3.55**)	5.62	(6.71***)	6.42	(7.21***)
−2 log-likelihood	1726.042		1236.153		1339.145	
Akaike's information criterion	1732.139		1239.347		1344.207	
Hurvich and Tsai's criterion	1732.203		1239.353		1344.283	
Bozdogan's criterion	1741.116		1246.028		1358.411	
Schwarz's Bayesian information criterion	1739.187		1243.125		1352.395	

*$p < .10$.
**$p < .05$.
***$p < .01$.

Notes: MCPd1 = increasing returns (dummy1), MCPd2 = constant returns (dummy2), and the base is decreasing returns to marketing communications expenditures; we do not show nonsignificant interaction terms. Bayesian information criterion imposes a severe penalty for more complicated model specifications. In the stock return model, ROA = unanticipated ROA (or time-series residuals), and MCP measures are all unanticipated residuals because of the efficient-markets rationale from the accounting literature. Because stock return is different from Tobin's q, we used the differences of MCP (or unanticipated MCP) for stock return model and used the levels of MCP for the Tobin's q model.

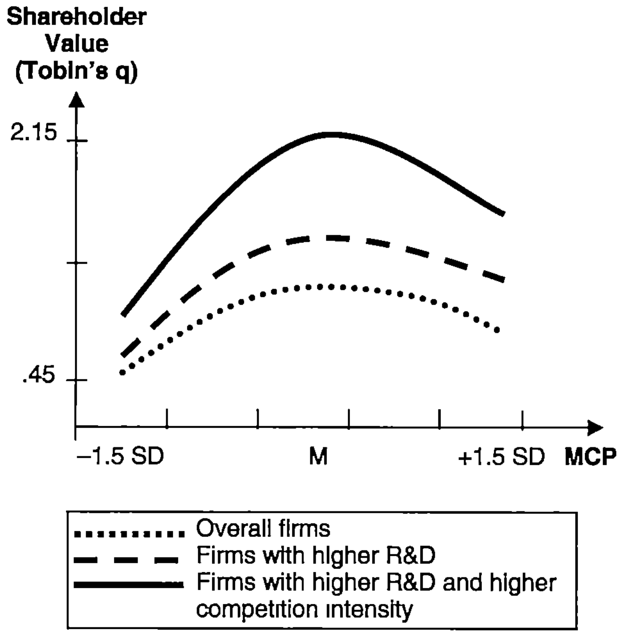
evidence of the complementary rents of marketing and R&D (Dutta, Narasimhan, and Rajiv 1999; Griffin and Hauser 1996; Gupta, Raj, and Wilemon 1987; Hayes, Wheelwright, and Clark 1988) but also supports the RBV that marketing and R&D resources should be configured as performance-enhancing bundles.

To test H_3, that MCP has the strongest influence on market value in firms that have higher R&D intensity *and* function in more competitive markets, we created three-way interaction terms among MCP dimensions, R&D, and competition intensity. As we show in Table 4, two of the three-way interactions are positive and significant (MCP1 × R&D × CPI: $t = 3.84$, $p < .01$; MCP3 × R&D × CPI: $t = 6.18$, $p < .01$), which indicates that a firm's Tobin's q is greater when the firm excels in MCP and R&D and operates in competitive markets over time. This finding suggests that in a competitive market, firms that have higher R&D and efficiently invest in their marketing communication programs are most likely to obtain sustainable competitive advantages dynamically (see Dutta, Narasimhan, and Rajiv 1999; Srivastava, Shervani, and Fahey 1998).

Because we hypothesized the impact of MCP as curvilinear relationships in H_1, we also include the product terms between the moderators (R&D and competition intensity) and the quadratic form of the MCP dimensions. As we show in Table 4, we find some interesting results. For example, because the linear × quadratic interaction terms (MCP1-squared × R&D and MCP2-squared × CPI: t = –2.27 and –3.36, respectively) are significant and negative, we conclude that overly high levels of MCP harm the firm's Tobin's q to a greater degree if the firm has greater R&D and/or operates in more competitive markets. This conclusion is also confirmed by the negative three-way linear × quadratic interaction term (MCP1-squared × R&D × CPI: t = –2.25). Overall, these findings are insightful; when firms face tougher competition and seek to capitalize on R&D, unbounded improvements in MCP (i.e., relentlessly cutting advertising and promotion programs) would be especially harmful to future performance. Thus, these findings lend further support to H_1, H_2, and H_3. Following the steps that Aiken and West (1991) recommend, we illustrate these curvilinear and moderated relationships among Tobin's q, MCP, R&D, and competition intensity in Figure 4.

However, in contrast to the strong support of the hypotheses in the Tobin's q model, the support in the stock return model is less significant. More specifically, in the stock return model, we used the unanticipated changes of ROA and MCP dimensions (Jacobson and Aaker 1987; Mizik and Jacobson 2003), because according to the efficient markets theory, the current stock price incorporates all previous information with profit implications. That is, abnormal stock returns in the financial markets react only to unanticipated information (e.g., Erickson and Jacobson 1992; Kothari 2001). In other words, stock prices are essentially a random (strong or weak) walk. Although criticism of such efficient market arguments exists, in general, the literature seems to support them (e.g., Chauvin and Hirschey 1993; Mizik and Jacobson 2003). Therefore, we define unanticipated changes of ROA and MCP as their deviations of the series from predictions based on prior information (a proxy for expectations). We operationalized these deviations as the residuals from a time-series forecast model (e.g., Griliches 1995; Mizik and Jacobson 2003). As a further clarification, because stock return is different from Tobin's q (Erickson and Jacobson 1992), we used the differences of MCP (or unanticipated MCP) for stock return model and the levels of MCP for the Tobin's q model in Equation 2. As we report in Table 4, overall, the results provide less support for the hypotheses in the stock return model. In particular, the results indicate that an unanticipated change in marketing managerial efficiency has a positive influence on forward-looking performance of stock return (MCP1: t = 4.03, $p < .01$), but this influence becomes negative (MCP1-squared: t = –2.21, $p < .01$) at too high increases in MCP1, as we expected. Thus, the data further support H_1 that MCP1 has an inverted U-shaped influence on stock return. However, only one linear × quadratic interaction term in stock return (MCP1-squared × CPI: t = –1.83, $p < .1$) is negative, as we expected, and none of the three-way linear × quadratic interaction terms are significant. Thus, these findings show much less support for H_2 and H_3 in the stock return model. These less significant results in the stock return model (compared to Tobin's q) are not totally unexpected; a recent study by Mittal and colleagues (2005) also finds that their marketing variables strongly influence Tobin's q but do not affect stock return significantly.

FIGURE 4
The Curvilinear Influences of MCP on Shareholder Value



We also examine the influences of MCP on backward-looking performance with ROA. As we report in Table 4, changes in MCP are less significantly related to ROA; only two dimensions are significant statistically, whereas four dimensions are related to shareholder value. These limited findings regarding the influence of MCP on ROA (MCP1: t = 3.81, $p < .01$; MCP2: t = 1.86, $p < .1$) and the impact of few quadratic terms are consistent with prior literature. Empirically, existing studies (e.g., Chauvin and Hirschey 1993; Erickson and Jacobson 1992) have established a positive link between the multiperiod accumulated effective advertising expenditures and profitability ROA. Intuitively, improving MCP by cutting costs increases ROA because cutting operating expenses should consistently enhance accounting profitability ratios (Aaker and Carman 1982).

Note that because the total effects of MCP include (1) direct main effects, (2) indirect effects through ROA, and (3) moderated effects with R&D and competition intensity on Tobin's q and stock return, the total impact may not strictly follow an inverted U shape. In particular, although total effects (direct, indirect, and moderated impact) of MCP have an inverted U shape for the Tobin's q model, as we expected, this is not the case for the stock return model. Nevertheless, because modern marketing theory (Rust et al. 2004) is more concerned with marketing productivity's implications for the forward-looking shareholder value and less concerned with implications for the backward-looking ROA profitability, interpretations of our results should focus more on direct main effects of MCP and their moderated effects with R&D and competition intensity (which have an inverted U shape, as hypothesized) than on the indirect effects through ROA, a backward-looking, short-term performance measure.

Additional Analyses

We adopted several steps to conduct additional analyses to ensure the robustness of our results. Because Malmquist modeling is nonparametric and may be sensitive to data extreme points and measurement error (Banker, Charnes, and Cooper 1984; Färe Grosskopf, and Lovell 1994; Färe et al. 2004), we repeat the modeling analyses with different combinations of variables (i.e., two outputs and four inputs, three outputs and three inputs, and different lags of advertising and R&D in adjusted sales revenue). The results pertaining to MCP across the different specifications are highly correlated (i.e., smallest $r = .83$, $p < .01$), and the longitudinal changes of MCP are robust. Furthermore, we checked whether our results were robust when using two-year-based sales growth (t versus t – 1). In doing so, we reran all the 39,872 + 22,784 linear programming models with the simple annual growth rate (from t – 1 to t). These results of different dimensions of MCP measures (MCP1, MCP2, MCP3) are highly consistent. The correlation results for the three pairs of MCP based on the three-year compounded sales growth and MCP based on the two-year sales growth were $r = .92$, .92, and .91 (all $ps < .01$). In addition, as we discussed previously, because there is a reverse-causality concern between financial performance and *Fortune*'s reputation measure, we conducted additional analyses. That is, we teased out the reverse causality of corporate reputation with the residual approach (Roberts and Dowling 2002). Then, we used the estimated "clean" residual corporate reputation (without reverse causality) and conducted one more round of estimation of all the 39,872 + 22,784 linear programming models. We found that the MCP results are robust. Across the dimensions, the correlations r for the three pairs of MCP based on the original score of corporate reputation and MCP based on the residual of corporate reputation are .89, .86, and .88 (all $ps < .01$). Table 5 reports a summary of the different sensitivity analyses in MCP results. Thus, it seems that our empirical work is reliable, and the MCP results are reasonably credible.

In addition, recall that MCP is a rather complex transformation of the three outputs that may correlate with MCP (i.e., significant correlations include $r_{(\text{MCP1 and sales level})} = .11$, $r_{(\text{MCP1 and sales growth})} = .14$, $r_{(\text{MCP1 and reputation})} = .12$, $r_{(\text{MCP2 and sales level})} = .12$, and $r_{(\text{MCP2 and reputation})} = .13$). Thus, we test whether a more parsimonious model might perform just as well. That is, we added the three outputs as covariates on top of those controls in Equation 2 and reestimated this more parsimonious model. The results on the impact of MCP still hold. In particular, the effects of MCP on Tobin's q are robust and in the expected inverted U shape for both MCP1 and MCP3, though MCP2-squared is still nonsignificant as in the base model ($p > .05$). Moreover, in an alternative approach, we tested our hypotheses with a hierarchical regression model (see Luo, Rindfleisch, and Tse 2007; Slotegraaf, Moorman, and Inman 2003) to account for the correlations between MCP and the three outputs and to examine the additional explanatory power of MCP on influencing Tobin's q (beyond that of the three outputs). In particular, we first regress MCP against the three outputs to obtain residuals that have parceled out the correlations with the three outputs. We then employ these residuals as MCP measures when reestimating Equation 2. Again, the results of this alternative modeling approach fail to

TABLE 5
Sensitivity Results of MCP

	Overall Productivity Change 1995–2002	Productivity Change 1995–1996	Productivity Change 1996–1997	Productivity Change 1997–1998	Productivity Change 1998–1999	Productivity Change 1999–2000	Productivity Change 2000–2001	Productivity Change 2001–2002
MCP overall (baseline)	1.282 (.823)	1.231 (.699)	1.213 (.754)	1.141 (.603)	1.223 (.604)	1.307 (.642)	1.652 (1.479)	1.207 (.480)
MCP overall (different sales growth measure)	1.279 (.908)	1.227 (.713)	1.221 (.675)	1.138 (.642)	1.219 (.610)	1.310 (.683)	1.646 (1.511)	1.209 (.486)
MCP overall (different corporate reputation measure)	1.212 (.910)	1.209 (.732)	1.224 (.862)	1.093 (.727)	1.218 (.715)	1.295 (.829)	1.663 (1.385)	1.192 (.523)

Notes: We based MCP baseline results on the three-year compounded sales growth rate and corporate reputation. We based MCP results with different sales growth measure on the two-year simple sales growth rate. We based MCP results with different corporate reputation measure on corporate reputation after accommodating the reverse-causality issue between financial performance and *Fortune* reputation measure.

reject our hypotheses of the inverted U shape and moderated impact of MCP on Tobin's q.

Furthermore, employing Hausman's (1978) tests, we compared the random effects model with the fixed effects models (firm-specific, year-specific fixed effects; see Arellano 2003; Boulding 1990; Boulding and Christen 2003). The Hausman test essentially tests the null hypothesis that the coefficients estimated by the relatively efficient random effects model are the same as the coefficients estimated by the relatively consistent fixed effects model (e.g., Baltagi 2001). A nonsignificant p value of this test would suggest that the use of the random effects model is appropriate, and a significant p value would indicate that the use of the fixed effects model is appropriate. As we report in Table 6, the Hausman test result is not significant ($\chi^2_{\text{diff}} = 2.788$, $p > .05$), thus favoring the specified random effects model in Equation 2. This is consistent with the similar findings in prior studies that also favored random effects modeling (e.g., Anderson, Fornell, and Mazvancheryl 2004; Mittal et al. 2005; Rao, Agarwal, and Dahlhoff 2004). Using the Bayesian information criterion, we conducted further likelihood ratio tests and found that the model with two-way random effects in Equation 2 outperformed ($p < .05$) the rival models with either time or firm random effects. In addition, because random effects modeling assumes an exchangeable covariance structure, we repeat the analyses with different specifications of covariance structure, including unstructured, AR(1), AR(1) heterogeneous, Toeplitz heterogeneous, and compound symmetry (Arellano 2003; Baltagi 2001; Boulding and Staelin 1993). As we show in Table 6, we fail to obtain significant differences (largest $\chi^2_{\text{diff}} = 3.081$, $p > .05$) across these covariance structure specifications.

Finally, in Equation 2, we specified random intercepts models (Gonul and Srinivasan 1993) in testing the direct and moderated impact of MCP on Tobin's q. However, because firms might differ in effectiveness in their deployment and exploitation of MCP, we also tried to test the robustness by fitting random coefficient models (Gonul and Srinivasan 1993; Murthi, Srinivasan, and Kalyanaram 1996; Rao, Agarwal, and Dahlhoff 2004). Using likelihood ratio tests, we failed to reject the null hypothesis that there is no significant variation in firm financial value due to the effects of random coefficients. Overall, we believe that the results are fairly robust.

Discussion

Marketing's credibility has been questioned, and the need to link marketing expenditures to financial performance has been stressed. However, the impact of MCP on firm market value has been unclear. In response to these issues, we developed a measure of MCP based on the neoclassical production theory and RBV. Our study tests a conceptual framework that accommodates both the curvilinear influence of MCP on shareholder value over time and the moderating roles of competition and R&D intensity. The key takeaway points of this study are that (1) firms that are more productive in converting advertising and promotion resources into marketing outputs may create greater shareholder value over time, (2) too great an increase in MCP can also be harmful, and (3) these results are especially robust for firms with intense R&D expenditures and that operate in competitive markets.

Implications for Marketing Researchers

Our study answers Rust and colleagues' (2004, p. 83) call for new methodology to measure marketing productivity, in which they clearly identify a key gap in the literature when they state that "few methods currently exist for comprehensively modeling the chain of marketing productivity all the way from tactical actions to financial impact or firm value." Our article explores this issue by introducing Malmquist modeling to measure MCP. The Malmquist approach has direct and handy implications for marketing modelers and strategists interested in evaluating the chain of marketing

TABLE 6
Random Effects Model Specifications Results

	Test Statistics	p Value	Conclusion
Model Specification Comparisons			
Random firm and time effects versus fixed firm and time effects	$\chi^2_{\text{diff}} = 2.788$	>.05	Random effects model is supported.
Random firm and time effects versus random firm effects only	$\chi^2_{\text{diff}} = 21.853$	<.05	Two-way random effects model is preferred.
Random firm and time effects versus random time effects only	$\chi^2_{\text{diff}} = 33.096$	<.05	Two-way random effects model is preferred.
Likelihood Ratio Test Comparing Different Covariance Structure in Random Effects Model			
Unstructured versus AR(1)	$\chi^2_{\text{diff}} = 2.675$	>.05	Estimates are robust across covariance structure specifications.
AR(1) versus AR(1) heterogeneous	$\chi^2_{\text{diff}} = 2.846$	>.05	Estimates are robust across covariance structure specifications.
AR(1) heterogeneous versus Toeplitz heterogeneous	$\chi^2_{\text{diff}} = 2.933$	>.05	Estimates are robust across covariance structure specifications.
Toeplitz heterogeneous versus compound symmetry	$\chi^2_{\text{diff}} = 3.081$	>.05	Estimates are robust across covariance structure specifications.

productivity and efficiency in novel ways. For example, this technique can be applied to benchmark consumer decision-making productivity and price promotion productivity with time-series consumer panel data (e.g., Dominick's [a grocery store] scanner data, *Consumer Reports*, ACNielsen data) and Internet pricing productivity and cross-channel efficiency dynamically with online clickstream data (e.g., comScore data, firm customer relationship management databases).

In addition, our Malmquist approach extends research on how to test the RBV empirically with scientific modeling techniques. Unlike most regression approaches that compare against average performers and neutralize firm differences with a central tendency, the Malmquist technique deals with best practices and above-average players. Similarly, the RBV is a theory about extraordinary performers (Hansen, Perry, and Reese 2004) and above-normal firms that outperform their industries (Wiggins and Ruefli 2005), a point with which Malmquist measurement modeling fits neatly.

The existing literature seems to be enthusiastic about improving MCP, but it ignores its potential dark side (with one notable exception being the work of Anderson, Fornell, and Rust [1997]). We suggest that an unrestricted increase of MCP may be harmful and cause negative market returns. Achieving overly high productivity may not be the most optimal way to improve firm market value, because doing so may hurt service quality and customer satisfaction, which are critical for firm future performance (Anderson, Fornell, and Rust 1997). The results of the negative quadratic influence of MCP may help explain the "lost" credibility of marketing; unbridled increases in MCP generate suboptimal market value for the firm over time, and only a well-balanced increase in MCP will produce higher Tobin's q and stock returns.

A burgeoning stream of research has begun to explore the linear, monotonic relationship between shareholder value and its predictive marketing variables (e.g., Anderson, Fornell, and Mazvancheryl 2004; Dutta, Narasimhan, and Rajiv 1999; Gruca and Rego 2005; Lee and Grewal 2004). Our findings of the nonlinear relationships between MCP and firm shareholder value extend this line of research. The financial literature similarly recommends such nonlinear models as the proper way to model stock price changes and stock returns (e.g., Hiemstra and Jones 1994). Therefore, we encourage marketing researchers to look beyond simple, linear relationships in determining firm market value.

The supported impact of interactions between MCP and R&D intensity adds more evidence to Dutta, Narasimhan, and Rajiv's (1999) argument about the complementary rents of marketing and R&D capabilities. Our finding of a positive influence in the three-way interaction among MCP, R&D, and competition intensity on Tobin's q also suggests a hidden point; namely, the synergy between marketing and R&D may be even greater when firms are motivated by market competition for Schumpeter-type entrepreneurial rents (Schumpeter 1934; Wiggins and Ruefli 2005). Other possible contextual moderators should be considered in future efforts to evaluate marketing productivity. In addition, it would be appealing for research to test our framework in a broader sense by including the influences of other perception-based factors, such as pricing strategies and market orientation. We were not able to do so empirically because of the limited availability of archival data. Thus, we call for further efforts with both primary survey data and secondary archival data to extend this line of research.

Implications for Marketing Practitioners

Where is marketing's credibility? Contrary to the common criticisms about advertising accountability and the decline of advertising and promotion productivity, our results indicate a positive influence of MCP on shareholder value. These findings help justify advertising and sales promotion spending, promote the creditability of the marketing function, and establish marketing as a distinct capability within the firm. To put the results into perspective, we calculate that for a typical firm in our sample, which has an average market value of approximately $43 billion, a 1% improvement in MCP (managerial efficiency) will add approximately $45 million more returns in the following year. For a high-tech firm such as Dell, with a market value of around $80 billion, a 1% increase in MCP translates into a substantial financial return: approximately $85 million.

Our Malmquist approach to MCP offers marketers a valuable tool for benchmarking and evaluating the results of advertising and promotion expenditures dynamically. To sell marketing and advertising programs to financial officers, marketers need to quantify the marketing budget and demonstrate marketing's contribution to shareholder value. In this context, our Malmquist-based scientific technique helps solve the thorny situations corporate marketers are currently facing:

> For many years, corporate marketers have walked into budget meetings like neighborhood junkies. They couldn't always justify how well they spend or what difference it all made. They just wanted more money. Old tricks simply don't work, there's hardly a marketing executive today who isn't demanding a more scientific approach to help defend marketing strategies in front of the chief financial officer. Marketers want to know the actual value of each dollar.... The push is from top ranks. [Chief executive officers, chief financial officers,] and even board directors have relentlessly cut costs in every corner of their companies.... The bean counters know that marketing matters. But they're hazy about how much or what kind. (*BusinessWeek* 2004, p. 112)

Although improving MCP largely promotes firm performance, blindly pushing for high MCP can be harmful for firm market value. The nonlinear findings of negative returns on MCP suggest that managers should resist the temptation to improve productivity simply by relentlessly cutting corners. Given a trade-off between productivity and customer satisfaction (Anderson, Fornell, and Rust 1997), an extremely lean marketing budget is not necessarily optimal. In addition, unbounded improvements in MCP, including cutting advertising and promotion programs, would be especially harmful for future performance if the firm faces tougher competition and seeks to capitalize on its R&D capabilities. Thus, these findings advise managers to take a cautious view of improving the productivity of their advertising and promotion expenditures. In conclusion, our study may help management achieve a more complete understanding of MCP, benchmark it against best practices over time, and examine its shareholder value in more depth and sophistication so that the marketing productivity crisis can be reversed.

Appendix
Malmquist Linear Programming for MCP

According to Malmquist (1953), a productivity change measure can be calculated by comparing the inputs of a firm at times t and t + 1 according to the maximum factor, by which the input in time t could be decreased such that the firm would continue to produce the same level of outputs at time t + 1. Caves, Christensen, and Diewert (1982) extend this idea and develop the total factor productivity change from t to t + 1; they label it "the Malmquist productivity measure." Combining Caves, Christensen, and Diewert's productivity index and Farrell's (1957) Pareto-based efficiency concept, Färe and colleagues (Färe et al. 1992; Färe, Grosskopf, and Lovell 1994; Färe et al. 2004) construct the linear-programming- (DEA-) based Malmquist productivity change from t to t + 1. This measure of MCP for each firm can be calculated with a series of linear programming models. The composite score of MCP for firm$_o$, or M_o, requires two single-period and two mixed-period efficiencies for the DEA programming (Banker, Charnes, and Cooper 1984; Charnes, Cooper, and Rhodes 1978), as follows:

$$\text{(A1)}\quad M_o = \left[\frac{D_o^t(x_o^{t+1}, y_o^{t+1})}{D_o^t(x_o^t, y_o^t)} \times \frac{D_o^{t+1}(x_o^{t+1}, y_o^{t+1})}{D_o^{t+1}(x_o^t, y_o^t)}\right]^{1/2},$$

where $D_o^t(x_o^t, y_o^t)$, $D_o^{t+1}(x_o^{t+1}, y_o^{t+1})$, $D_o^{t+1}(x_o^t, y_o^t)$, and $D_o^t(x_o^{t+1}, y_o^{t+1})$ are relative efficiencies estimated by four linear programming models.

After some mathematical manipulations, we can decompose the Malmquist measurement (M_o) of MCP into several components (e.g., Caves, Christensen, and Diewert 1982; Färe, Grosskopf, and Lovell 1994). As we show subsequently, the first component of MCP1 is the change of productivity from time t to t + 1 that is due to managerial efficiency (or the effects of moving toward the best-practice frontiers). This efficiency represents how much improvement in MCP can be achieved if executives enhance the skills of managing their marketing communication compared with the best-practices. The second component, MCP2, is change in productivity from time t to t + 1 that is due to capital efficiency (or the effects of moving between the best-practice frontiers with different returns-to-scale model specifications). This efficiency represents how much improvement in MCP can be achieved if executives promote the economies of scale when using the marketing communication capitals. The third dimension of MCP3 is the change of productivity from time t to t + 1 that is due to best-practice shift efficiency (or the effects of moving away from the best-practice frontiers from t to t + 1). This efficiency represents how much improvement in MCP can be achieved through factors that are not related to marketing communication managerial skills but rather to uncontrollable factors such as industry communication creativity innovation and environment. Finally, the returns to scale of marketing communications capital (or the effects of moving along the best-practice frontiers) is a categorical variable that can increase, be constant, or decrease in returns to scale (Caves, Christensen, and Diewert 1982; Färe, Grosskopf, and Lovell 1994). The equation is as follows:

$$\text{(A2)}\quad M_o(t, t+1) = \underbrace{\frac{D^{VRS}_{o,t+1}(x_o^{t+1}, y_o^{t+1})}{D^{VRS}_{o,t}(x_o^t, y_o^t)}}_{\text{MCP1}}$$

$$\times \underbrace{\left[\frac{D^{CRS}_{o,t+1}(x_o^{t+1}, y_o^{t+1})/D^{VRS}_{o,t+1}(x_o^{t+1}, y_o^{t+1})}{D^{CRS}_{o,t}(x_o^t, y_o^t)/D^{VRS}_{o,t}(x_o^t, y_o^t)}\right]}_{\text{MCP2}}$$

$$\times \underbrace{\left[\frac{D^{CRS}_{o,t}(x_o^{t+1}, y_0^{t+1})}{D^{CRS}_{o,t+1}(x_o^{t+1}, y_o^{t+1})} \times \frac{D^{CRS}_{o,t}(x_o^t, x_o^t)}{D^{CRS}_{o,t+1}(x_o^t, y_o^t)}\right]^{1/2}}_{\text{MCP3}}$$

For the sake of brevity, we illustrate the details of only one linear programming model for $D_o^{t+1}(x_o^t, y_o^t)$. The other linear programming models are similar, with only some changes in the superscript of either t or t + 1. The term $D_o^{t+1}(x_o^t, y_o^t)$ assesses the amount by which marketing communication inputs at time t can be reduced but produce the same amount of output levels at time t when firm M_o is compared with the best-practice frontier at time t + 1.

$$D_o^{t+1}(x_o^t, y_o^t) = \min \theta_o,$$

subject to (1) multiple outputs constraints:

$$\sum_{j=1}^{n} \lambda_j SAL_j^{t+1} \ge SAL_o^t,$$

$$\sum_{j=1}^{n} \lambda_j SAG_j^{t+1} \ge SAG_o^t, \text{ and}$$

$$\sum_{j=1}^{n} \lambda_j REP_j^{t+1} \ge REP_o^t;$$

(2) multiple inputs constraints:

$$\sum_{j=1}^{n} \lambda_j ADB_j^{t+1} \le \theta\ ADB_o^t,$$

$$\sum_{j=1}^{n} \lambda_j ADP_j^{t+1} \le \theta\ ADP_o^t,$$

$$\sum_{j=1}^{n} \lambda_j ADO_j^{t+1} \le \theta\ ADO_o^t, \text{ and}$$

$$\sum_{j=1}^{n} \lambda_j SAP_j^{t+1} \le \theta\ SAP_o^t;$$

and (3) returns-to-scale constraints for variable returns-to-scale models:

$$\sum_{j=1}^{n} \lambda_j = 1, \text{ and}$$

$$\lambda_j \geq 0, j = 1, 2, \ldots, j = n;$$

where $\lambda_j \geq 0$; j = 1, 2, ..., n = sample size; r = 1, 2, ..., s outputs; and i = 1, 2, ..., m inputs. There are four marketing expenditure inputs (firms' expenditure in broadcast [ADB], print [ADP], outdoor [ADO], and sales promotion [SAP]) and three marketing outputs (sales level [SAL], sales growth rate [SLG], and corporate reputation [REP]). Adjusted sales capture the advertising carryover effect and the influence of other variables, such as R&D. The parameter theta essentially means the degree to which the inputs can be reduced without hurting the levels of outputs (e.g., Horsky and Nelson 1996; Luo 2004; Murthi, Srinivasan, and Kalyanaram 1996). Thus, the optimal minimized theta represents the best scenario in which a firm can achieve the highest possible efficiencies in marketing communication overall productivity benchmarked against the best-practice frontier. In addition, $D_0^t(x_0^t, y_0^t)$ measures the amount by which marketing communication inputs at time t can be reduced while producing the same amount of output levels at time t compared with best practices at time t; $D_0^{t+1}(x_0^{t+1}, y_0^{t+1})$ measures the amount by which marketing communication inputs at time t + 1 can be reduced while producing the same amount of output levels at time t + 1 compared with best practices at time t + 1; and $D_0^t(x_0^{t+1}, y_0^{t+1})$ measures the amount by which marketing communication inputs at time t + 1 can be reduced while producing the same amount of output levels at time t + 1 compared with best practices at time t.

REFERENCES

Aaker, David and James A. Carman (1982), "Are You Over Advertising?" *Journal of Advertising Research*, 22 (4), 57–70.

——— and Robert Jacobson (1994), "The Financial Information Content of Perceived Quality," *Journal of Marketing Research*, 31 (May), 191–201.

——— and ——— (2001), "The Value Relevance of Brand Attitude in High-Technology Markets," *Journal of Marketing Research*, 38 (November), 485–93.

Achenbaum, Alvin (1992), "Reversing the Advertising Productivity Crisis," *Marketing Management*, 1 (3), 22–27.

Aiken, Leona and Stephen G. West (1991), *Multiple Regression: Testing and Interpreting Interactions*. Thousand Oaks, CA: Sage Publications.

Ambler, Tim (2004), "Differing Dimensions," *Marketing Research*, 16 (Fall), 8–13.

Anderson, Eugene W., Claes Fornell, and Sanal K. Mazvancheryl (2004), "Customer Satisfaction and Shareholder Value," *Journal of Marketing*, 68 (October), 172–85.

———, ———, and Roland T. Rust (1997), "Customer Satisfaction, Productivity, and Profitability: Differences Between Goods and Services," *Marketing Science*, 16 (2), 129–45.

Arellano, Manuel (2003), *Panel Data Econometrics*. New York: Oxford University Press.

Baltagi, Badi (2001), *Econometric Analysis of Panel Data*. New York: John Wiley & Sons.

Banker, R., A. Charnes, and W.W. Cooper (1984), "Some Models for Estimating Technical and Scales Inefficiencies in Data Envelopment Analysis," *Management Science*, 30 (9), 1078–1092.

Barney, Jay B. (1986), "Strategic Factor Markets: Expectations, Luck, and Business Strategy," *Management Science*, 32 (10), 1231–41.

Boulding, William (1990), "Commentary on 'Unobservable Effects and Business Performance: Do Fixed Effects Really Matter?'" *Marketing Science*, 9 (Winter), 88–91.

——— and Markus Christen (2003), "Sustainable Pioneering Advantage? Profit Implications of Market Entry Order," *Marketing Science*, 22 (3), 371–92.

——— and Richard Staelin (1993), "A Look on the Cost Side: Market Share and Competitive Environment," *Marketing Science*, 12 (Spring), 144–66.

Bucklin, Louis P. (1978), *Productivity in Marketing*. Chicago: American Marketing Association.

BusinessWeek (2004), "Making Marketing Measure Up," (December 13), 112–14.

Byrd, John and Kent Hickman (1992), "Do Outside Directors Monitor Managers?" *Journal of Financial Economics*, 32 (2), 195–221.

Caves, D.W., L.R. Christensen, and W.E. Diewert (1982), "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity," *Econometrica*, 50 (6), 1414–39.

Charnes, Abe, William W. Cooper, and E. Rhodes (1978), "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research*, 3 (4), 429–44.

Chauvin, Keith W. and Mark Hirschey (1993), "Advertising, R&D Expenditures, and the Market Value of the Firm," *Financial Management*, 22 (Winter), 128–40.

Chen, Yao and Agha Ali (2004), "DEA Malmquist Productivity Measure: New Insights with an Application to Computer Industry," *European Journal of Operational Research*, 159 (1), 239–49.

Dekimpe, M. and Michael Hanssens (1995), "The Persistence of Marketing Effects on Sales," *Marketing Science*, 14 (1), 1–21.

Dittus, E.C. and M. Kopp (1990), "Advertising Accountability in the 1990s: Moving from Guesswork and Gut Feelings," *Journal of Advertising Research*, 30 (6), RC7–RC12.

Dutta, Shantanu, Om Narasimhan, and Surendra Rajiv (1999), "Success in High-Technology Markets: Is Marketing Capability Critical?" *Marketing Science*, 18 (4), 547–68.

Erickson, Gary and Robert Jacobson (1992), "Gaining Comparative Advantage Through Discretionary Expenditures: The Returns to R&D and Advertising," *Management Science*, 38 (September), 1264–79.

Färe, R., S. Grosskopf, B. Lindgren, and P. Roos (1992), "Productivity Change in Swedish Pharmacies 1980–1989: A Non-Parametric Malmquist Approach," *Journal of Productivity Analysis*, 3 (2), 85–102.

———, ———, and C.A.K. Lovell (1994), *Production Frontiers*. Cambridge, UK: Cambridge University Press.

———, ———, Barry Seldon, and Victor Tremblay (2004), "Advertising Efficiency and the Choice of Media Mix: A Case of Beer," *International Journal of Industrial Organization*, 22 (4), 503–522.

Farrell, M.J. (1957), "The Measurement of Productive Efficiency," *Journal of the Royal Statistical Society*, 120 (3), 253–81.

Fischer, Harald and Timothy G. Pollock (2004), "Effects of Social Capital and Power on Surviving Transformational Change: The Case of Initial Public Offerings," *Academy of Management Journal*, 47 (4), 463–81.

Fortune (2005), "America's Most Admired Companies," (March 7), 68.

Gonul, Fusun and K. Srinivasan (1993), "Modeling Unobserved Heterogeneity in Multinominal Logit Models: Methodological and Managerial Implications," *Marketing Science*, 12 (3), 213–29.

Griffin, Abbie and J.R. Hauser (1996), "Integrating R&D and Marketing: A Review and Analysis of the Literature," *Journal of Product Innovation Management*, 13 (3), 191–215.

Griliches, Zvi (1995), "R&D and Productivity," in *Handbook of the Economics of Innovation and Technological Change*, Paul Stoneman, ed. Oxford: Blackwell Publishers, 52–89.

——— and Jerry A. Hausman (1986), "Errors in Variables in Panel Data," *Journal of Econometrica*, 31 (1), 93–118.

Gruca, Thomas S. and Lopo L. Rego (2005), "Customer Satisfaction, Cash Flow, and Shareholder Value," *Journal of Marketing*, 69 (August), 115-130.

Gupta, A.K., S.P. Raj, and D. Wilemon (1987), "Managing the R&D-Marketing Interface," *Research Management*, 30 (2), 38–43.

Hansen, Mark H., Lee Perry, and C. Shane Reese (2004), "A Bayesian Operationalization of the Resource-Based View," *Strategic Management Journal*, 25 (13), 1279–95.

Hausman, Jerry A. (1978), "Specification Tests in Econometrics," *Econometrica*, 46 (November), 1251–71.

Hayes, R.H., S.C. Wheelwright, and K.N. Clark (1988), *Dynamic Manufacturing*. New York: The Free Press.

Hiemstra, Craig and Jonathan Jones (1994), "Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation," *Journal of Finance*, 49 (December), 1639–64.

Holstius, K. (1990), "Sales Response to Advertising," *International Journal of Advertising*, 9 (1), 38–56.

Horsky, Dan and Paul Nelson (1996), "Evaluation of Salesforce Size and Productivity Through Efficient Frontier Benchmarking," *Marketing Science*, 15 (4), 301–320.

Houston, Mark and Shane Johnson (2000), "Buyer–Supplier Contracts Versus Joint Ventures: Determinants and Consequences of Transaction Structure," *Journal of Marketing Research*, 37 (February), 1–15.

Jacobson, Roberts (1990), "Unobservable Effects and Business Performance: Reply to the Comments of Boulding and Buzzell," *Marketing Science*, 9 (1), 92–95.

——— and David A. Aaker (1987), "The Strategic Role of Product Quality," *Journal of Marketing*, 51 (October), 31–44.

Joseph, Kissan and Vernon J. Richardson (2002), "Free Cash Flow, Agency Costs, and the Affordability Method of Advertising Budgeting," *Journal of Marketing*, 66 (January), 94–107.

Keller, Kevin (1998), *Strategic Brand Management*. Upper Saddle River, NJ: Prentice Hall.

——— and David Aaker (1993), "The Effects of Sequential Introduction of Brand Extensions," *Journal of Marketing Research*, 29 (February), 35–50.

Kim, John, Mukesh Bhargava, and Sridhar Ramaswami (2001), "Advertising Productivity: Developing an Agenda for Research," *International Journal of Advertising*, 20 (4), 431–54.

Kothari, S.P. (2001), "Capital Markets Research in Accounting," *Journal of Accounting and Economics*, 31 (1–3), 105–231.

Kumar, V. and J. Andrew Petersen (2004), "Maximizing ROI or Profitability," *Marketing Research*, 16 (Fall), 28–34.

Lee, Ruby and Rajdeep Grewal (2004), "Strategic Responses to New Technologies and Their Impact on Firm Performance," *Journal of Marketing*, 68 (October), 157–71.

Luo, Xueming (2004), "Data Envelopment Analysis: A Management Science Tool for Scientific Marketing Research," *Journal of Marketing Research*, 42 (August), 113–16.

——— and Naveen Donthu (2001), "Benchmarking Advertising Efficiency," *Journal of Advertising Research*, 41 (6), 7–18.

——— and ——— (2005), "Assessing Advertising Media Spending Inefficiency in Generating Sales," *Journal of Business Research*, 58 (1), 28–36.

———, Aric Rindfleisch, and David Tse (2007), "Working with Rivals: The Impact of Competitor Alliances on Financial Returns to Competitor-Oriented Firms," *Journal of Marketing Research*, 44, forthcoming.

Malmquist, S. (1953), "Malmquist Index Numbers and Indifference Surfaces," *Trabajos de Estadistica*, 4 (3), 209–242.

McAllister, P.H. and D. McMannus (1993), "Resolving the Scale Efficiency Puzzle in Banking," *Journal of Banking and Finance*, 17 (2–3), 389–405.

McGuire, Jean, Thomas Schneewels, and Ben Branch (1990), "Perceptions of Firm Quality: A Cause or Result of Firm Performance," *Journal of Management*, 16 (1), 167–80.

McLaughlim, John (1997), "Accountability Is New Name of the Ad Game," *Marketing News*, 31 (4), 5.

Mittal, Vikas, Eugene Anderson, Akin Sayrak, and Pandu Tadikamalla (2005), "Dual Emphasis and the Long-Term Financial Impact of Customer Satisfaction," *Marketing Science*, 24 (4), 544–55.

Mizik, Natalie and Robert Jacobson (2003), "Trading Off Between Value Creation and Value Appropriation: The Financial Implications of Shifts in Strategic Emphasis," *Journal of Marketing*, 67 (January), 63–76.

Mueller, Dennis (1990), "The Persistence of Profits in the United States," in *The Dynamics of Company Profits*, Dennis Mueller, ed. Cambridge, UK: Cambridge University Press, 35–37.

Murthi, B.P.S., Kannan Srinivasan, and Gurumurthy Kalyanaram (1996), "Controlling for Observed and Unobserved Managerial Skills in Determining First-Mover Market Share Advantages," *Journal of Marketing Research*, 33 (August), 329–37.

Ng, Christina (2005), "An Empirical Study on the Relationship Between Ownership and Performance in a Family-Based Corporate Environment," *Journal of Accounting, Auditing and Finance*, 20 (Spring), 121–46.

Penrose, Edith T. (1959), *The Theory of the Growth of the Firm*. Oxford: Basil Blackwell.

Prasad, Arbind and Subrata Sen (1999), "Are Firms Advertising Too Much," paper presented at the Marketing Science Conference, Syracuse University (May).

Rao, Vithala, Manoj K. Agarwal, and Denise Dahlhoff (2004), "How Is Manifest Branding Strategy Related to the Intangible Value of a Corporation," *Journal of Marketing*, 68 (October), 126–41.

Reichheld, Frednerick F. (1996), *The Loyalty Effect*. Boston: Harvard Business School Press.

Roberts, Peter and Grahame Dowling (2002), "Corporate Reputation and Sustained Superior Financial Performance," *Strategic Management Journal*, 23 (2), 1077–1093.

Rust, Roland (1986), *Advertising Media Models: A Practical Guide*. Lexington, MA: Lexington Books.

———, Tim Ambler, Gregory Carpenter, V. Kumar, and Rajendra K. Srivastava (2004), "Measuring Marketing Productivity: Current Knowledge and Future Directions," *Journal of Marketing*, 68 (October), 76–89.

———, Christine Moorman, and Peter R. Dickson (2002), "Getting Return on Quality: Cost Reduction, Revenue Expansion, or Both?" *Journal of Marketing*, 66 (October), 7–24.

Schumpeter, J.A. (1934), *The Theory of Economic Development*. Cambridge, MA: Harvard University Press.

Sevin, C.H. (1965), *Marketing Productivity Analysis*. New York: McGraw-Hill.

Slotegraaf, Rebecca J., Christine Moorman, and J. Jeffrey Inman (2003), "The Role of Firm Resources in Returns to Market Deployment," *Journal of Marketing Research*, 40 (August), 295–309.

Sougiannis, Theodore (1994), "The Accounting Based Valuation of Corporate R&D," *The Accounting Review*, 69 (1), 44–68.

Srivastava, Rajendra, Tasadduq Shervani, and Liam Fahey (1998), "Market-Based Assets and Shareholder Value: A Framework for Analysis," *Journal of Marketing*, 62 (January), 2–18.

———, ———, and ——— (1999), "Marketing, Business Processes, and Shareholder Value: An Organizationally Embedded View of Marketing Activities and the Discipline of Marketing," *Journal of Marketing*, 63 (Special Issue), 168–79.

Strong, Edward (1925), "Theories of Selling," *Journal of Applied Psychology*, 9 (February), 75–86.

Szymanski, David M., Sundar G. Bharadwaj, and P Rajan Varadarajan (1993), "An Analysis of the Market Share–Profitability Relationship," *Journal of Marketing*, 57 (July), 1–18.

Vakratsas, Demetrios and Tim Ambler (1999), "How Advertising Works: What Do We Really Know?" *Journal of Marketing*, 63 (January), 26–43.

Wernerfelt, Birger (1984), "A Resource-Based View of the Firm," *Strategic Management Journal*, 5 (2), 171–80.

Wheelock, David and Paul Wilson (1999), "Technical Progress, Inefficiency, and Productivity Change in U.S. Banking, 1984-1993," *Journal of Money, Credit, and Banking*, 31 (2), 212–34.

Wiggins, Timothy and Robert R. Ruefli (2005), "Schumpeter's Ghost: Is Hypercompetition Making the Best of Times Shorter?" *Strategic Management Journal*, 26 (9), 887–911.

Now in Its Third Year of
Bringing You the Best in Market Research

Review of Marketing Research

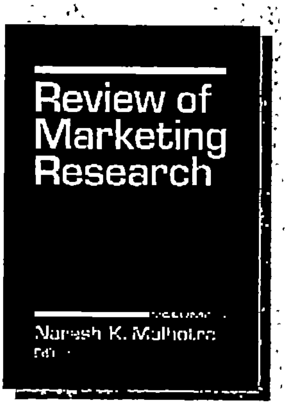
Edited by **Naresh K. Malhotra**, Georgia Institute of Technology

Review of Marketing Research provides current, state-of-the-art articles by the marketing field's leading researchers and academicians. Unlike other research publications that impose rigid limitations on the length of articles, *RMR* publishes longer contributions that are not only theoretically rigorous, but also offer richer detail, including literature reviews, cutting-edge methodologies, empirical studies, emerging trends, international developments, guidelines for implementation, and suggestions for future theory development and testing.

"Fills a major gap in the marketing discipline. The editorial board consists of some of the top scholars in the field. ... I expect that *RMR* will play a significant role in shaping the future of research conducted in marketing."
– Vijay Mahajan, The University of Texas at Austin

"Clearly the series will serve to promote high quality research in the marketing discipline by informing all of us of the present state of the art. ... I recommend that every serious marketing scholar, marketing researcher, and library should subscribe to the review series."
– Gary L. Frazier, University of Southern California

Volume 3, October 2006
232 pages ISBN 0-7656-1306-9 / ISSN 1548-6435 Hardcover $99 95



All volumes of *Review of Marketing Research* are available online at no additional cost.

For complete information on any volume in the series, visit www.mesharpe.com/rmr.htm

M.E. Sharpe

TO ORDER: Call 800-541-6563 or 914-273-1800 • Fax 914-273-2106 • www.mesharpe.com
AD612X

Erica Mina Okada

Upgrades and New Purchases

In many industries, firms introduce successive generations of new and enhanced product models to remain competitive. Product enhancements target first-time purchasers and consumers who already own an existing product and are considering upgrading to a new and enhanced product. Upgrade decisions differ from new purchase decisions because they are hindered by the psychological costs associated with the purchase price spent on the existing product. This research examines how the enhanced product can be positioned relative to the existing product to mitigate the psychological costs and facilitate upgrade purchases. A series of studies demonstrate that consumers with existing products are more likely to upgrade when the enhanced product is generally dissimilar to the existing product. Upgraders have stronger preferences than first-time purchasers for enhanced products with new features rather than for improvements on existing features, and they prefer enhancements that focus on a few key features rather than a general enhancement of all features.

Continual product enhancement is a key component of a firm's success (Urban and Hauser 1993). In many industries, the market size typically does not expand to accommodate the proliferation of increasingly sophisticated models, so the commercial success of each generation of product enhancements depends to a large extent on current owners of the existing model who upgrade to the new model before the useful life of the existing model expires. When Sony introduced the PlayStation 2 (PS2) in March 2000, the highly popular original PlayStation that it launched only four years earlier had an installed base of 30 million in the United States. By the end of 2003, PS2 had an installed based of almost 30 million (Peterson 2004). Given the total size of the U.S. market, PS2 sales must have come not only from new buyers in the market but also significantly from owners of the original PlayStation who upgraded to PS2.

Product enhancements target both new purchasers and "upgraders"—that is, consumers who already own a product and upgrade to a new, enhanced product—but upgrade decisions are different from new purchase decisions. The upgrader already has an existing product, so the marginal benefit of the enhanced product will likely not be as great as for a new buyer who does not have an existing product. Another economics-based explanation is that the upgrader may be more knowledgeable about the product than a new purchaser, and his or her preferences may have evolved differently over usage. There are also the psychological factors based on mental accounting (see Thaler 1980, 1999). Okada (2001) demonstrates how upgrade decisions are hindered by the psychological costs associated with taking an existing product out of commission while it is still functional. This is a psychological cost for upgraders but not for first-time purchasers.

This research builds on existing literature on mental accounting (Thaler 1980, 1999) and sunk costs (Arkes 1996; Arkes and Blumer 1985; Heath 1995; Heath and Fennema 1996; Heath and Soll 1996; Okada 2001) to explore how those psychological costs can be mitigated in ways that are within a marketer's control. The psychological costs become a lesser impediment to upgrading when the consumer perceives the enhanced product as dissimilar to the existing product. Therefore, between two enhanced products that may be equally attractive to a new buyer with no prior purchase, the upgrader will find the one that is more dissimilar to the existing product to be relatively more attractive. Even after accounting for any differences in knowledge and preferences (see Joshi and Sharma 2004) between upgraders and new purchasers, the differences still maintain. To appeal to upgraders, then, marketers should position their enhanced product as being dissimilar to the existing product.

Marketers can enhance a product by improving existing product attributes and by adding new attributes. On the basis of research on attribute alignability (Markman and Gentner 1993; Markman and Medin 1995), I define the former as "alignable enhancements" and the latter as "nonalignable enhancements." Research on perceived similarity (Tversky 1977) and product comparability (Johnson 1984) suggests that alignable enhancements would make the new product more similar to the existing product, and nonalignable enhancements would make the new product more dissimilar. Compared with a first-time purchase, nonalignable enhancements would be relatively more desirable in an upgrade than proportional alignable enhancements. Furthermore, enhancements that focus on a select few features would be relatively more desirable in an upgrade than more general enhancements across all features.

Theoretical Development

In an upgrade, a consumer who already owns a product that is still functional purchases a new enhanced product, which makes the existing product redundant (Okada 2001).

Erica Mina Okada is an assistant professor, University of Washington Business School (e-mail: emokada@u.washington.edu). The author gratefully acknowledges Teck Ho, Douglas MacLachlan, and Richard Yalch for their helpful comments and Joel Berg for his help in data collection. She also thanks the three anonymous *JM* reviewers.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Although consumers who purchase an enhanced product may not necessarily dispose of the existing product, this conceptualization of "upgrade" is consistent with the colloquial usage of the word. A traveler who upgrades opts for an enhanced class of travel instead of traveling coach. A user who upgrades to a new computer may keep the existing one as a backup but will likely use the new, enhanced computer primarily. Thus, I contrast upgrade decisions to new purchase decisions.

Upgrading to Enhanced Products That Are Dissimilar Versus Similar

People keep a mental account of the costs and benefits of transactions over time (Gourville and Soman 1998; Heath and Fennema 1996; Thaler 1980, 1985, 1999). When a consumer purchases a product, he or she opens a mental account and posts the purchase price as a negative entry into the account, but the consumer feels no pain at the time of payment, because he or she psychologically links the costs and benefits of the transaction (Hirst, Joyce, and Schadewald 1994; Kahneman and Tversky 1984). As the consumer uses the product, the benefits from consumption are posted in the same account as positive entries. As the cumulative enjoyment from consumption increases, the consumer gets his or her money's worth from the purchase. The account is closed when he or she finishes consuming the product.

If an upgrader purchases an enhanced product, he or she will no longer use the existing product, which triggers the closing of that mental account. There is a psychological cost associated with closing the existing account in the red, that is, before the consumer has gotten his or her money's worth from the existing product. That is what hinders people from upgrading, even when the marginal cost–benefit comparison favors it (Okada 2001).

In mental accounting terms, what differentiates upgrade decisions from new purchase decisions is the "write-off" of the existing product. The purchase price spent on the existing product is a sunk cost (see Arkes 1996; Arkes and Blumer 1985), which normatively should be ignored in prospective decisions, but the motivation to recoup the sunk cost by continuing to use the existing product hinders consumers' upgrade decisions (Okada 2001). I propose that the sunk cost in the existing product is more likely to be disassociated from the upgrade decision and, therefore, to become less of an impediment when the enhanced product is dissimilar to the existing product.

Categorization is the basis for mental accounting (Thaler 1985). People tend to mentally account for similar products together and for dissimilar products separately (Henderson and Peterson 1992), and prospective purchase decisions are influenced more by prior purchase decisions in similar categories (Heath and Soll 1996). For example, a dinner out at a restaurant and a theater ticket are more alike than a pair of slacks and a theater ticket. Thus, a consumer who may not buy a theater ticket if he or she has already gone out to a dinner earlier in the week—because that would be overspending in the entertainment category—may still buy the same theater ticket after having bought a pair of slacks that are comparable in price to the dinner out. The prospective theater purchase is affected more by a prior restaurant dinner purchase because the two are alike. The prospective clothing purchase is affected less because it is dissimilar to the restaurant dinner.

Likewise, in an upgrade, the decision to purchase an enhanced product should be affected less by the existing product the more it is dissimilar to it. When the enhanced product is more dissimilar to the existing product, the prospective write-off of the sunk cost in the existing product should become more disassociated from the consideration of the enhanced product, and it should become less of an impediment to the upgrade decision. Thus, upgraders should have relatively stronger preferences than new purchasers for enhanced products that are more dissimilar to their existing products.

> H_1: When a person has paid for and already owns an existing product, an enhanced product that is dissimilar to the existing one should be relatively more preferable than a proportionally enhanced product that is similar.

In the current context, "proportional" means the extent to which a new purchaser would consider two enhanced products comparable in their overall desirability, if it were not for the existing product. For example, a person may normally deem a six-day ski vacation to Aspen and a six-day cruise to Acapulco to be equivalently desirable. However, if he or she first paid a nonrefundable deposit on a two-day ski trip to the Poconos and was later presented with the opportunity to go on either of the more "enhanced" vacations instead, he or she may be more likely to choose the more dissimilar Acapulco cruise than the more similar Aspen ski trip. Presumably, the psychological cost of writing off the nonrefundable deposit would be more likely to be disassociated from the decision to upgrade to the cruise than to a better ski trip. The current theory centers on the link between the (dis)similarity of the enhanced product and the extent to which the sunk cost in the existing product is considered in, and thus impedes, the upgrade decision.

Positioning New Products as Nonalignable Versus Alignable Enhancements

If dissimilar enhanced products are more desirable to upgraders, the products should be positioned as such to target them. Two items will be perceived as more similar if they share more attributes in common and more dissimilar if they are characterized by attributes that are unique to only one (Johnson 1984; Tversky 1977). Attributes that two (or more) products share in common are alignable attributes, and those that are possessed only by one are nonalignable attributes (Markman and Medin 1995). Products with alignable attributes tend to belong to similar product categories (Markman and Gentner 1993), but (dis)similarity is a relative measure. Even in the same general product category, some item pairs may be more similar than other pairs. A Mazda Miata roadster and BMW Z3 roadster are more similar than a Mazda Miata and a Lexus RX330 SUV. The former pair has relatively more alignable attributes between them than the latter pair.

In terms of a multiattribute model that explains consumers' preferences and choices as a function of products' attribute levels and importance (see Bettman, Capon, and Lutz 1975; Green, Carroll, and Goldberg 1981; Kahn and Meyer 1991; Shocker and Srinivasan 1979), an enhanced product, relative to an existing product, is characterized by a combination of higher levels of existing attributes and the addition of new attributes.

The term "alignable enhancements" refers to the improvement of existing attributes, and "nonalignable enhancements" refers to the addition of new attributes. For example, Panasonic introduced the Versio model of the mobile phone in 2000; it weighed 85 grams and had no browser capabilities. The new X70 model weighs 10% less, which is an alignable enhancement; has browsing capabilities, which is a nonalignable enhancement; and is an overall enhancement of the Versio. Perceived similarity between pairs of items increases with the number of features that the two share in common and decreases with the number of features that are distinctive to just one (Tversky 1977). In light of research on similarity (Tversky 1977) and product comparability (Johnson 1984), nonalignable enhancements should make the enhanced product more dissimilar to the existing product than should proportional alignable enhancements.

H_1 proposed that consumers should find it easier to upgrade to enhanced products that are dissimilar to the existing product than to products that are similar to it. If nonalignable enhancements make the new product more dissimilar, compared with new purchasers, upgraders should have relatively stronger preferences for nonalignable enhancements than for proportional alignable enhancements.

> H_2: When a person has paid for and already owns an existing product, an enhanced product that is characterized by additional attributes should become relatively more desirable than a proportionally enhanced product that is characterized by improvements to existing attributes.

In a previous article (Okada 2001), I examined upgrades characterized by alignable enhancements and found that the mental cost of writing off the existing product directly impeded the consumer's upgrade decision. The current research complements my previous work by differentiating between nonalignable and alignable enhancements and by showing how this differentiation alters the effect. The mental cost of upgrading should be mitigated when the enhancements are nonalignable.

Adding features makes the enhanced product more dissimilar to the existing one, but it may not always be feasible or practical. This research compares consumers' perceptions of proportional product enhancements. However, when consumers are unfamiliar with the added features, they may not value the enhancements as much (Zhang and Fitzsimons 1999; Zhang and Markman 1998). Furthermore, if consumers perceive added features as trivial, the enhancements may backfire and actually diminish the value of the product (Mukherjee and Hoyer 2001; Simonson, Carmon, and O'Curry 1994). An enhanced product can still be positioned as dissimilar from the existing product without adding new features and based on just the existing attributes by focusing the enhancements on a select few attributes.

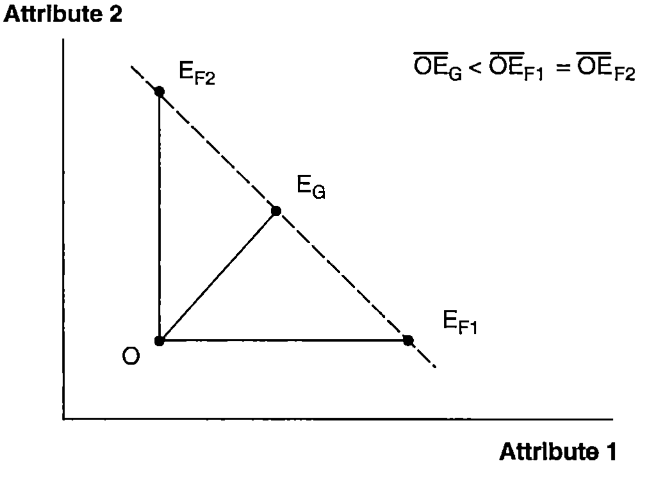
Based on a multiattribute model, it is possible to achieve proportional levels of product enhancement by higher increases in the levels of a select few attributes or by lesser increases in the levels of all attributes. I define the former approach as "focused enhancements" and the latter as "general enhancements." Focused enhancements tend to make the new product more dissimilar to the existing product.

When product-market structure is represented spatially, product attributes define the dimensions of space. Consumers perceive two products that are closer together as more similar (Best 1976; Ghose 1998; Glazer and Nakamoto 1991). Figure 1 represents a simplified, two-dimensional product space defined by two equally important attributes (Attributes 1 and 2) and an existing product situated at O. The axes are attribute ratings. The existing product can be enhanced proportionally by improving the rating of Attribute 1 to E_{F1} or by improving the rating of Attribute 2 to E_{F2}. Anywhere along the segment that connects E_{F1} to E_{F2} represents the same proportional level of enhancement. It can likewise be achieved by enhancing both Attributes 1 and 2 to E_G. A general enhancement of both attributes, such as E_G, results in a closer distance from the existing product O and relatively greater similarity to O. When the enhancements are focused on either one or the other of the attributes, such as E_{F1} or E_{F2}, a greater distance from O results and, thus, a relatively greater dissimilarity to O.

The principle generalizes to products that are characterized by more than two attributes. Enhancements that are focused on a select few attributes result in greater distances from and more dissimilarity to the existing product than do general enhancements on all attributes.

> H_3: When a person has paid for and already owns an existing product, an enhanced product that is characterized by focused enhancements in a select few attributes should be more desirable than a proportionally enhanced product that is characterized by more general enhancements across all attributes.

FIGURE 1
Enhancements to Maximize Distance



In this research, I explore the addition of attributes and the improvement of existing attributes as the primary ways to enhance products, but I do not explore overall enhancements in which existing attributes are taken away. When any significant part of the existing product is deleted in an overall enhancement, the enhanced product would be an additional purchase rather than an upgrade. For example, when the microwave oven was introduced, it may have been an overall enhancement compared with the conventional oven, but many of the functionalities of a conventional oven were deleted. Furthermore, most people who owned conventional ovens did not upgrade from conventional to microwave ovens and discontinue using their conventional ovens. Rather, they typically bought the microwave ovens in addition to the conventional ovens. An upgrade effectively terminates the consumer's usage of the existing product. A series of four studies examine how the sunk cost in an existing product can be mitigated to facilitate upgrade decisions.

Study 1

People are less willing to upgrade to an enhanced alternative when there is a sunk cost in the existing alternative. Study 1 was designed to gain insight into whether the dissimilarity of the enhanced product mitigates the negative effect of the sunk cost and, in turn, facilitates upgrade decisions. Study 1 addresses the question, Is it easier to upgrade to enhanced alternatives that are dissimilar to the existing alternative than to proportionally desirable enhanced products that are similar?

Method

Ninety-four undergraduate students participated in a paper-and-pencil study. Data were collected over two sessions, six weeks apart. There were a total of four items in this study: two moderate and two enhanced. The two moderate items were described as a ticket to a Seattle Mariners playoff game in Safeco Field in balcony seats behind left field and a ticket to a rock concert by their favorite band in the Key Arena with end seats to the left of the stage in the top row. The two enhanced items were a ticket to a Seattle Mariners playoff game in Safeco Field in front-row seats behind the first-base line and a ticket to a rock concert by their favorite band in the Key Arena with front-row balcony seats and a backstage pass. Participants were randomly assigned to one of four different combinations of one moderate item and one enhanced item.

In one session, participants were asked to indicate their willingness to pay (WTP) for their randomly assigned pair of one moderate item and one enhanced item. Participants' WTP in this session measured the desirability of the alternatives in the absence of an existing product that would be made redundant by the purchase.

Six weeks later, participants were given scenarios in which they made plans to go to a moderate event, after which an opportunity arose to go to a more desirable, enhanced event. They were asked how much they would be willing to pay to go to the enhanced event. Participants were presented with the same combination of moderate and enhanced events in these scenarios as in the first session. The difference in the second session was that the moderate alternative was presented as the existing alternative, and the enhanced alternative was made available as a new alternative after the existing plans were made. For half of the participants, the existing alternative was the moderate baseball ticket, and for the other half, it was the moderate concert ticket. In all cases, the enhanced event occurred at the same time as the existing event, so participants could go to either one or the other event but not to both. Furthermore, the ticket for the existing alternative was nonrefundable and nontransferable; therefore, purchasing the new enhanced alternative would render the existing ticket worthless. This created an upgrade situation. For half of the participants, the enhanced alternative was similar to the existing alternative; that is, an enhanced baseball ticket became available when there was an existing moderate baseball ticket, or an enhanced concert ticket became available when there was an existing moderate concert ticket. For the other half, the enhanced alternative was dissimilar to the existing alternative; that is, an enhanced concert ticket became available when there was an existing moderate baseball ticket, or an enhanced baseball ticket became available when there was an existing moderate concert ticket.[1]

There were two levels of sunk cost to manipulate how much participants had paid for the existing ticket. Half of the participants had paid $15 for their existing moderate alternatives. The other half first paid $15 but then got an unexpected refund of the purchase price for winning a drawing, so they ended up with no money sunk into the ticket (see Soman and Cheema 2001). If participants were to infer the desirability of the moderate alternative on the basis of whether they had chosen to pay for it, the inferred value should be the same in both conditions. Their WTP for the enhanced event was the dependent measure.

The study was a 2 (type of existing alternative: moderate baseball or moderate concert) × 2 (similar or dissimilar enhanced product) × 2 (sunk cost or no sunk cost) design. Participants' WTP in the first session measured their valuation of the enhanced items in situations that more closely resembled new purchases. In the second session, participants indicated their WTP for the same products but in an upgrade context.

Results and Discussion

The data collected in the first session served as a manipulation check of the stimuli for the existing and enhanced alternatives and confirmed that the average WTP for the moderate baseball and moderate concert tickets (M = $24.11 and $27.64, respectively) was lower than the average WTP for the enhanced baseball and enhanced concert tickets (M = $51.35 and $50.98, respectively). On average, the enhanced baseball and enhanced concert tickets were comparably desirable to each other and were more desirable than the existing moderate baseball and moderate concert ticket alternatives, which were also comparably desirable to each other. This suggests that the scenarios were fair representations of upgrade situations in which people who already owned a product had an opportunity to acquire an enhanced alternative that would render the preexisting alternative redundant. I also used these data to calculate the marginal benefit of the enhanced product compared with the moderate product (the difference in the WTP of the former less the latter). Overall, the average marginal benefit of the enhanced product over the moderate was $25.30.

---

[1] A manipulation check of the dissimilar and similar conditions was performed on 25 separate participants. All 25 participants indicated that the enhanced baseball ticket was more similar to the moderate baseball ticket than to the enhanced concert ticket. All 25 participants also indicated that the enhanced concert ticket was more similar to the moderate concert ticket than to the enhanced baseball ticket.

The data collected in the second session suggest that participants were willing to pay $22.41 for the enhanced event when there was a preexisting moderate event for which they already had tickets. The results appear in Table 1.

When the enhanced product was similar to the existing product, participants were willing to pay significantly less ($t = 3.4$, $p < .01$) for the enhanced alternative when they had sunk costs in the existing product (M = $7.70) than when there was no sunk cost (M = $30.40). This is consistent with the literature on sunk costs. However, when the enhanced product was dissimilar to the existing product, there was no difference ($t = .35$, $p = .72$) between the WTP for the enhanced product when there were sunk costs (M = $23.64) and the WTP for the enhanced product when there was no sunk cost (M = $27.08).

To manipulate the level of sunk cost and equalize any inferred desirability of the moderate alternative on the basis of whether participants had paid for it, the no-sunk-cost condition was presented as a situation in which respondents had first chosen to pay $15 but were unexpectedly refunded. This inevitably created a condition of a windfall gain in the no-sunk-cost condition. The literature suggests that people treat windfall gains differently from money they earn or already have (O'Curry and Strahilevitz 2001), but this does not offer an alternative explanation for the current findings, because it was in the sunk-cost group, in which there was no windfall gain, that the differences in WTP between the dissimilar and the similar enhancement conditions were significant.

An analysis of covariance ($F_{5, 88} = 32.57$, $p < .001$) of the WTP for the enhanced product, on the (dis)similarity of the enhanced alternative to the existing alternative, on whether there was a sunk cost in the existing alternative, on the interaction between the two aforementioned variables, on whether the existing alternative was a moderate baseball or concert ticket, and on the marginal benefit of the enhanced alternative compared with the existing alternative as a covariate, demonstrated that WTP for the enhanced alternative was significantly lower ($F_{1, 88} = 4.16$, $p = .04$) when there was a sunk cost in the existing alternative. This is consistent with the findings in previous research (Okada 2001); that is, presumably, the sunk cost in the existing product impedes upgrade decisions. Furthermore, the negative effect of the sunk cost on WTP for the enhanced alternative was significantly mitigated ($F_{1, 88} = 8.38$, $p < .01$) when the enhanced alternative was dissimilar to the existing product, in support of H_1. The covariate was also significant; intuitively, respondents who perceived a higher marginal benefit ($F_{1, 88} = 139.86$, $p < .001$) in the enhanced product were willing to pay a higher price. In terms of the baseball and concert events being well matched in desirability, there was no difference ($F_{1, 88} = .80$, $p = .37$) whether the existing alternative was the moderate baseball or concert ticket.

When people already own an existing alternative, and an opportunity arises to upgrade to a superior alternative, they are willing to pay more for the new, enhanced alternative if it is dissimilar to the existing alternative. If the new alternative is for the same event but in better seating, the WTP is significantly lower, presumably because the sunk cost in the existing alternative hinders the upgrade decision. This is consistent with previous research on the sunk-cost effect (Arkes 1996; Arkes and Blumer 1985; Heath 1995) and, more specifically, on product upgrade decisions (Okada 2001). Conversely, if the new alternative is dissimilar to the existing alternative, WTP is significantly higher and comparable to the marginal benefit, presumably because the sunk cost in the existing alternative becomes less relevant in considering the transaction. When the new alternative is dissimilar, it becomes easier to disassociate the write-off of the sunk cost, and the upgrade decision becomes essentially like a new purchase.

Study 2

The results of Study 1 support H_1 and demonstrate that people are willing to pay more to upgrade when the enhanced alternative is dissimilar to the existing alternative. The theoretical explanation is based on the differential effects of the sunk cost; that is, the write-off of the sunk cost becomes more disassociated with the upgrade decision when the enhanced alternative is more dissimilar to the existing product, which in turn facilitates upgrade decisions. Study 2 was designed to gain insight into the process and to measure the extent to which sunk costs influence

TABLE 1
Mean WTP for Similar Versus Dissimilar Upgrades in Conditions of Sunk Cost and No Sunk Cost

	Dissimilar Enhanced Product		Similar Enhanced Product		
	Baseball	Concert	Baseball	Concert	Total
No sunk cost	$26.15	$28.18	$31.15	$29.58	$28.78
Sunk cost	$25.00	$22.27	$ 7.58	$ 7.82	$15.49
Total	$25.43		$19.52		$22.41

upgrade decisions differentially when the enhanced product is dissimilar versus similar.

Method

One hundred forty-seven undergraduate students participated in this study, in which upgrade scenarios were presented using the same product stimuli as those in Study 1. Half of the participants had purchased moderate baseball tickets, and the other half had purchased moderate concert tickets as the original plan. Then an enhanced alternative became available after the existing plans were made. As before, both the enhanced and the moderate alternatives occurred on the same day, so participants could go to one or the other event but not to both. Furthermore, the ticket for the existing alternative was nonrefundable and nontransferable. The moderate and enhanced alternatives were described in the same way as in Study 1.

Whereas in Study 1, the sunk cost was manipulated, in this study, the sunk cost was $15 for all participants. The objective was to measure the differential effect of the same level of sunk cost on the upgrade decision as a function of the (dis)similarity of the enhanced product to the existing one.

As in Study 1, for half of the participants, the enhanced alternative was similar to the existing one; that is, they upgraded from the moderate baseball ticket to the enhanced baseball ticket or from the moderate concert ticket to the enhanced concert ticket. For the other half, the enhanced alternative was dissimilar; that is, they upgraded from the moderate baseball ticket to the enhanced concert ticket or from the moderate concert ticket to the enhanced baseball ticket.

In contrast to Study 1, which measured WTP, Study 2 offered the enhanced product for $20, and the dependent measure was whether participants chose to upgrade. This price was only $5 more than the $15 they paid for the moderate alternative, but the scenario was that a student agency at school was offering the enhanced tickets at a special student price, and the offer price did not directly reflect the market price. Participants also indicated the extent to which they would consider the $15 they had spent on the existing alternative in their decisions using a seven-point scale (0 = "the $15 would not be a consideration at all," and 6 = "the $15 would be an extremely significant consideration"). The prediction is that the $15 would be a lesser consideration when upgrading to a dissimilar enhanced product, which in turn should facilitate the upgrade decision.

Results and Discussion

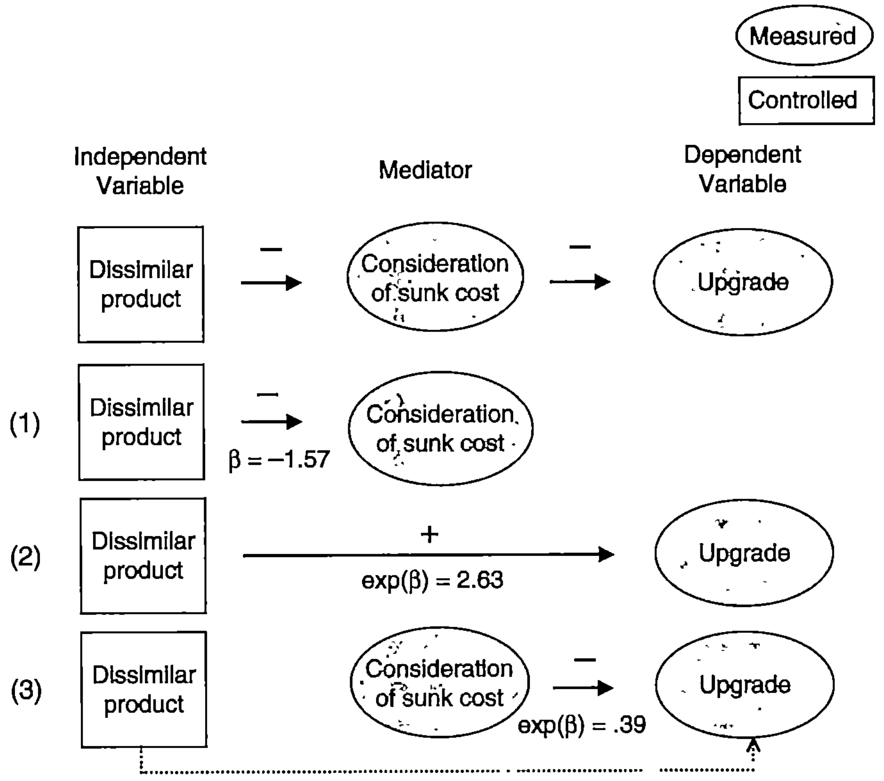
Overall, 87 of the 147 participants (59.2%) indicated that they would upgrade to the enhanced alternative. As in Study 1, whether the existing ticket was the moderate baseball or the moderate concert ticket made no difference (Z = .14) in the percentage of those who upgraded, so I aggregated the data across the two existing alternative conditions for the analysis and compared upgrade scenarios in which the enhanced alternative was dissimilar versus similar to the existing alternative.

Of the 77 participants in the dissimilar upgrade condition, 54 (70.1%) indicated that they would upgrade to the enhanced alternative for $20. This was significantly greater (Z = 2.83) than the 33 of 70 (47.1%) in the similar upgrade condition who indicated that they would upgrade. Furthermore, the dissimilar upgrade group (M = 2.1) considered the sunk cost in the existing ticket to a significantly lesser extent (t = –5.77, $p < .001$) than the similar upgrade group (M = 3.6). To measure the role of sunk cost, I asked participants directly about the extent to which they considered the $15; this method may arguably make the issue more salient. However, the sunk-cost measure was taken after the upgrade decision. Furthermore, the data indicate a greater extent of consideration of sunk cost in the similar than in the dissimilar group, but there is no theoretical reason that any collateral increase in salience of sunk cost should be different between the two groups.

I conducted a mediation analysis (Baron and Kenny 1986) to test the theory that when the enhanced product is dissimilar to the existing product, the sunk cost in the existing product will be considered to a lesser extent, which in turn will increase the likelihood of upgrading, as illustrated in Figure 2. First, the dissimilarity of the enhanced product significantly decreased the extent to which people considered sunk costs in their decisions (t = –5.77, $p < .001$; $\beta = -1.57$). Second, by a logistic regression, dissimilarity also significantly increased the likelihood of upgrading ($p < .01$). The Wald statistic was 7.85, and the effect on the odds ratio (exp[β]) was 2.63. Numbers greater than 1 indicate a positive effect on the odds ratio. Third, when the likelihood of upgrading was regressed against both the dissimilarity of the enhanced product and the consideration of sunk costs, again by logistic regression, only the sunk cost was significant ($p < .001$), and dissimilarity became not significant ($p = .62$). The Wald statistics were 34.19 and .25, respectively, and the respective effects on the odds ratio (exp[β]) were .39 and .80. Numbers less than 1 indicate a negative effect on the odds ratio, and this showed that upgrading became less likely when the sunk cost was considered to a greater extent.

One limitation that the design may impose is that because the price of the enhanced alternative was only $5 more than what participants had paid for the existing moderate alternative, they may have made inferences about the quality of the enhanced product, which would affect the likelihood of upgrading. The same results could be obtained if such negative-quality inferences are made to a greater extent for similar enhanced products than for dissimilar ones. However, participants were given descriptions of the moderate and enhanced alternatives in the same way as in Study 1; they were not just given the price. Furthermore, the enhanced alternatives were made available through a student agency at school at a special student price, and the offered price did not directly reflect the market price. Consistent with the proposed theory, the mediation analysis in Study 2 showed that participants considered the sunk cost in the existing product to a lesser extent when the enhanced product was dissimilar, which in turn facilitated the upgrade decision.

FIGURE 2
Mediation Analysis



Study 3

According to theories of similarity (Tversky 1979) and product comparability (Johnson 1984), adding new features should make the new product more dissimilar to the original product than proportionally improving existing features. Study 3 examines whether people are more likely to upgrade to a new product that is characterized by nonalignable enhancements or to upgrade to a new product that is characterized by proportional alignable enhancements.

Another objective of Study 3 is to determine whether the findings of Study 1 are more broadly generalizable. Because the concept of upgrading is relevant for durable-type products as well, Study 3 examined mobile phones. In addition, in Study 3, I collected data from consumers who actually purchased mobile phones, so the existing product in Study 3 is actual, not hypothetical.

Method

The 179 participants of this study were recruited at a local shopping center. The product stimulus was a Sony Ericsson T68i mobile phone. Participants were shown an actual model of the phone and were given a brief written description of the phone. Then, they completed a brief paper-and-pencil questionnaire and indicated how much they would be willing to pay for the new Sony Ericsson mobile phone. To make purchase scenarios viable and realistic and to control for knowledge and experience resulting from product usage, only the people who currently owned mobile phones were recruited. People participated on a volunteer basis, so the surveys were kept as succinct as possible.

The survey first asked how much they would be willing to pay for the new phone under one of two conditions: the upgrade or the new purchase condition. In the new purchase condition, participants were asked to imagine that they lost their existing phones. This created a situation in which there would be no existing phone to become obsolete as a result of the new purchase, and thus the phone would resemble a new purchase. The upgrade condition made no qualifications on participants' existing phones, and they were simply asked how much they would be willing to pay for the new phone. Because participants all had mobile phones, their existing phones would presumably become redundant if they purchased the new phone. In both the new purchase and the upgrade conditions, participants were asked to assume that there were no additional costs, such as termination or activation fees, and that the monthly charges would remain equivalent for an equivalent level of usage.

In the second part of the survey, participants compared the new phone with the mobile phones they actually owned in terms of the alignable and nonalignable attributes. The survey briefly defined alignable features as those that the two phones shared in common and nonalignable features as those that were unique to one phone. Participants indicated on a seven-point scale (–3 = "strongly disagree," and +3 = "strongly agree") whether the new phone was superior to their own phones in the alignable features and nonalignable features, respectively. The order between the two questions was randomized. In essence, Study 3 measured responses to hypothetical situations but from consumers considering their actual purchase history.

Results and Discussion

Of the 179 people who participated in the survey, 90 were in the new purchase condition, and 89 were in the upgrade condition. Respondents were willing to pay \$65.12 on average for the new phone, and WTP was higher ($t = 2.31$, $p = .02$) for the new purchase group ($M = \$75.98$) than for the upgrade group ($M = \$54.15$), which is intuitive because the people in the new purchase condition effectively had no working phone, and thus the new phone's marginal benefit would presumably be greater.

On average, participants perceived the new phone as superior in its alignable features ($M = .55$; $t = 4.74$, $p < .01$) and its nonalignable features ($M = 1.07$; $t = 7.13$, $p < .01$) compared with the phones that they owned. Therefore, the new phone was a reasonably enhanced phone. In terms of the random assignment of the participants to the new purchase and upgrade conditions, there were no differences between the two groups in the perceived levels of the new phone's alignable enhancements ($M = .62$ versus $.48$; $t = .60$, $p = .55$) or nonalignable enhancements ($M = 1.11$ versus 1.03; $t = .26$, $p = .80$).

H_2 predicted that in terms of upgrades, nonalignable enhancements would have a greater positive impact on purchase decisions than proportional alignable enhancements. Therefore, nonalignable enhancements should have a relatively greater positive effect on WTP in the upgrade condition, and the perceived levels of alignable enhancements should have a relatively greater positive effect on WTP in the new purchase condition.

I ran a regression to analyze the differential effects of alignable versus nonalignable enhancements in the upgrade versus new purchase conditions. The results appear in Table 2. First, I mean-centered the alignable and nonalignable enhancement measures by converting them to Z-scores, arriving at the alignable and nonalignable Z-scores. Second, to control for the multicollinearity between the alignable and the nonalignable measures ($\rho = .61$), I regressed the nonalignable Z-scores against the alignable Z-scores and calculated the residual of the nonalignable Z-scores after subtracting the component that is predicted by alignable Z-scores to arrive at the residual nonalignable Z-scores. The main regression analysis used these calculated variables.

The dependent variable was the WTP for the enhanced phone, and the explanatory variables were the alignable Z-score, the residual nonalignable Z-score, whether the phone was considered a new purchase or an upgrade, the interaction between the alignable Z-score and whether it was a new purchase or an upgrade, and the interaction between the residual nonalignable Z-score and whether it was a new purchase or an upgrade. The model had an R-square of .24. People were willing to pay less when the new phone was an upgrade to replace an existing phone than when it was essentially a new purchase ($t = -2.50$, $p = .01$). This is intuitive because the marginal benefit would be greater for those who do not have functional phones. People were willing to pay more for the new phone when they perceived it as superior to their current phones on the alignable attributes ($t = 3.52$, $p = .001$). Because there was a relatively high correlation between the alignable and the nonalignable enhancement measures, this significant effect captured the increased WTP due to nonalignable enhancements as well. So the residual nonalignable Z-score did not have a significant effect on WTP ($t = -1.22$, $p = .22$). However, in the upgrade condition, people were willing to pay more for nonalignable enhancements ($t = 4.20$, $p < .001$) but only marginally more for alignable enhancements ($t = -1.95$, $p = .05$). Even by a conservative test that measured the level of nonalignable enhancements only as a residual after subtracting the component that correlated with alignable enhancements, the analysis still demonstrated a differentially positive effect of nonalignable enhancements on WTP in upgrading.

Thus, Study 3 suggests that in upgrading, nonalignable enhancements are more desirable than proportional alignable enhancements. Presumably, the addition of new attributes makes the enhanced product more dissimilar, and the sunk cost in the existing product becomes less salient.

Study 4

Study 4 was designed to test the prediction that focused enhancements of a select few of the existing attributes would be more attractive to upgraders than proportional general enhancements of all existing attributes.

TABLE 2
Regression of WTP for New Mobile Phone

Variables	β	SE	t	*p*	VIF
Constant**	75.04	5.96	12.59	< .001	—
Alignable Z-score**	19.07	5.42	3.52	= .001	1.64
Residual nonalignable Z-score	–12.09	9.91	–1.22	= .22	3.43
Upgrade or new purchase*	–21.1	8.45	–2.50	= .01	1.00
Upgrade × alignable Z-score	–16.98	8.70	–1.95	= .05	1.64
Upgrade × residual nonalignable Z-score**	49.44	11.78	4.20	< .001	3.43

*Significant at $\alpha = .05$.
**Significant at $\alpha = .01$
Notes: VIF = variance inflation factor

Method

Data were collected at a dental clinic from patients who came in for their routine, semiannual checkups. They participated by completing short anonymous surveys on a volunteer basis. Patients who were scheduled for more invasive dental procedures were not included in the study. On arriving at the clinic for their routine checkups, patients were asked by the receptionist if they would be willing to fill out a short survey while they waited. The patients who volunteered to participate completed their surveys and returned them anonymously into a box.

The survey presented scenarios in which participants were going on a vacation over an upcoming long weekend and had made arrangements to stay at a moderately rated bed-and-breakfast. Then, a few days later, they found two other alternatives for accommodation, both more desirable than the one at which they made reservations and both priced at $80. They were asked to indicate which of the two more desirable accommodations they would choose or to indicate "neither" if they wanted to stay at the moderate bed-and-breakfast they found first. This was the dependent measure.

The moderate lodging and the two more desirable hotels were all rated on five different dimensions and were given one overall rating. The rating was on a five-point scale, anchored by "fair" (1) and "excellent" (5). Those comparisons appear in Table 3. The surveys presented the moderate lodging information first on the front page. The two more desirable alternatives with higher ratings were presented on the next page side by side. With one enhanced hotel, the enhancements were focused on fewer features, and with the other, the enhancements were more general.

There were two different levels of commitment to the moderate bed-and-breakfast. In the sunk-cost condition, participants paid a nonrefundable $25 deposit, which would "go to waste" if they chose to stay elsewhere. In the no-sunk-cost condition, they made a deposit, but the deposit was fully refundable. The prediction is that in the sunk-cost condition, which more closely resembles an upgrade decision, the focused enhancements would be relatively more attractive, and in the no-sunk cost condition, which more closely resembles a new purchase, the general enhancements would be relatively more attractive.

A separate manipulation check on a group of 150 participants measured the relative desirability of the two enhanced alternatives and their relative similarity to the moderate alternative. In a side-by-side choice between the two enhanced alternatives, there was no significant difference ($Z = 1.30$, $p = .10$) between the percentages of those who chose the focused enhancement as being more desirable ($n = 83$, 55.3%) and those who chose the general enhancement as being more desirable ($n = 67$, 44.7%). This suggested that the two enhanced alternatives were comparably desirable and helped rule out an alternative sunk-cost-based explanation, namely, that the prospect of having to write off sunk costs creates negative emotions, which in turn compromises the level of scrutiny in the decision making, and the choice ends up different and less optimal than if there was no sunk cost and there were no negative emotions. However, the results suggest that there was no clearly "optimal" choice from which participants would deviate with a lesser level of scrutiny due to any negative emotion.

In the second part of the manipulation check, the moderate alternative was introduced as a comparison, and participants were asked which of the two enhanced alternatives was more similar to the moderate alternative. Significantly, more people ($Z = 7.02$, $p < .001$) perceived the general enhancement ($n = 118$, 78.7%) than the focused enhancement ($n = 32$, 21.3%) as more similar to the moderate alternative. This confirmed the manipulation in the main study and was consistent with the premise for the hypothesis that the enhanced product would be perceived as more dissimilar to the moderate product with focused enhancements and more similar with general enhancements.

Results and Discussion

One hundred sixty-eight people participated in the main study over a four-week period. Overall, 109 (64.9%) chose the more generally enhanced alternative, and the other 59 (35.1%) chose the more desirable, focused-enhancement alternative. No one chose "neither." This indicated an overall preference for the generally enhanced alternative ($Z = 3.86$, $p < .001$). Although these results did not replicate the manipulation check results, which suggested no significant difference in desirability between the two alternatives, preferences were measured in a different context in the main study. In the manipulation check, just the two enhanced alternatives were presented and compared, without the moderate alternative. In the main study, the moderate product was presented first, and the two enhanced products were presented next, so they were presumably each evaluated relative to the moderate alternative. On the basis of decreas-

TABLE 3
A Comparison of the Moderate and Enhanced Alternatives in Study 3

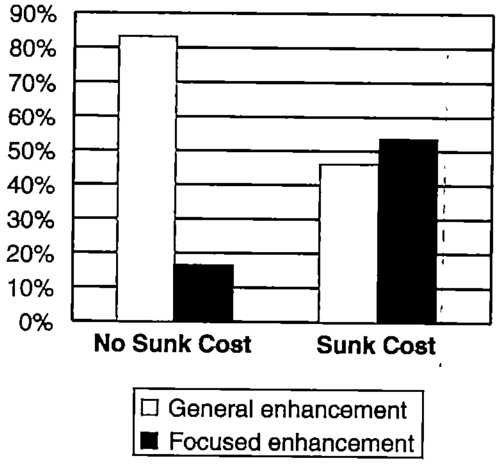
	Moderate Alternative	General Enhancement	Focused Enhancement
Location and access	2	3	5
Comfort and cleanliness of room	2	3	5
Friendliness of staff	2	3	2
Quality of food	2	3	2
Amenities	2	3	2
Overall rating	Moderate 2	Good 3	Good 3

ing marginal sensitivity in gains (Kahneman and Tversky 1979), smaller incremental gains across more attributes, which characterize the generally enhanced alternative, would have a greater positive impact on value than larger incremental gains across fewer attributes, which characterize the focused enhancement.

Eighty-four participants were in the no-sunk-cost group, and 84 participants were in the sunk-cost group. The results appear in Figure 3. In the no-sunk-cost group, 70 (83.3%) participants chose the more generally enhanced alternative, showing a strong preference ($Z = 6.11$, $p < .001$) for the general enhancement. In the sunk-cost group, the direction reversed, and 45 (53.6%) participants chose the more focused enhancement. The absolute preference for one alternative over the other may be driven by differences in the importance of the attributes. However, the objective of Study 3 was to determine whether there are any differences between upgraders and new purchasers in their preferences for focused versus general enhancements. Thus, I compared the preference for the focused enhancement with the preference for the general enhancement between the two groups. Consistent with H_3, the relative preference for the focused enhancement was greater in the sunk-cost group than in the no-sunk-cost group ($Z = 5.14$, $p < .001$).

The results of this study suggest that different types of product enhancements appeal to upgraders versus new purchasers. When enhancing an existing product on the basis of the existing attributes, focusing the enhancements on a few key attributes tends to be relatively more attractive to upgraders than to new purchasers. Presumably, focused enhancements result in products that are perceived as more dissimilar to the original product than proportional general enhancements, which makes it easier for upgraders to disassociate the write-off of the original product from the upgrade purchase decision.

FIGURE 3
Focused Versus General Enhancements With and Without Sunk Costs



General Discussion

It is helpful for marketers to understand how upgrade decisions differ from new purchases because many new products they introduce to the market target both new buyers and upgraders. Previous research has demonstrated that upgrade decisions are hindered by the mental costs associated with putting a still-functioning product out of commission, and consumers may forgo upgrades even when the marginal cost–benefit comparison is favorable. The current research suggests that depending on how the new product is positioned relative to the existing one, the sunk-cost effect can be mitigated, and the upgrade decision can be facilitated. Consumers have an easier time ignoring the sunk costs when upgrading to new products that are dissimilar to the existing product. Marketers can position their new products to be dissimilar by adding new features rather than improving existing ones and by focusing the enhancements on a few key attributes.

This research examined product enhancements by adding new attributes, but product enhancements can also combine both the addition of some new features and the deletion of some existing ones. Adding two new attributes and deleting one may make the enhanced product more dissimilar to the existing product than just adding one new attribute, assuming equally important attributes. According to the dissimilarity argument proposed here, the former may be relatively more desirable to the upgrader. However, loss aversion predicts that upgraders would respond unfavorably to deleted attributes. This research did not examine the interaction between adding new attributes and deleting existing ones. This is a direction for a future extension.

I considered upgrades in one-time consumable services (events, hotel accommodations) and durable-type products (phones). However, I did not compare one-time consumable products with repeat-usage products, and I did not consider any differential negative impact of sunk costs in upgrading. The prediction would be that the sunk cost may be a greater impediment for upgrading in durable-type products because of the repeat nature of the expected usage. With durable-type products, preferences for different types of upgrades may also change with the passage of time. For example, the desirability of new attributes may strengthen as consumers satiate with the existing attributes over time. Alternatively, it may weaken over time as consumers adapt to the existing attributes. These are also areas for further research.

The basic concept that consumers forgo purchasing a superior alternative to save previous expenditures from "going to waste" can also apply to products offered with loyalty programs. Suppose that only connecting flights on an existing airline were available out of a traveler's hometown airport. If a new airline enters the market and offers direct nonstop flight service, it would become the enhanced alternative. However, if travelers have accrued points with the existing airline that remain unused, these points, which they "paid for" by flying the connecting flights, would have to be "written off" if the travelers switched to the enhanced service. In addition to any economic reason for continuing to fly connecting flights, the mental cost of upgrading may also hinder frequent flyers from upgrading to the more convenient service.

Having spent money on an existing product that is still functional, upgraders tend to prefer different types of enhanced products compared with new purchasers. It is not just the differential knowledge and experience from ownership and usage that drive this pattern. Simply refunding the money spent (Studies 1 and 4) or rendering the existing product not functional (Study 3) changed people's preferences for the enhanced product. There is a mental cost to upgrading, and this research explores how the positioning of the enhanced product relative to the existing one can help mitigate that cost and facilitate upgrade decisions.

REFERENCES

Arkes, Hal R. (1996), "The Psychology of Waste," *Journal of Behavioral Decision Making*, 9 (3), 213–24.

——— and Catherine Blumer (1985), "The Psychology of Sunk Cost," *Organizational Behavior and Human Decision Processes*, 35 (February), 124–40.

Baron, Reuben M. and David Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (6), 1173–82.

Best, Roger J. (1976), "The Predictive Aspects of a Joint-Space Theory of Stochastic Choice," *Journal of Marketing Research*, 13 (May), 198–204.

Bettman, James R., Noel Capon, and Richard J. Lutz (1975), "Cognitive Algebra in Multi-Attribute Attitude Models," *Journal of Marketing Research*, 12 (May), 151–64.

Ghose, Sanjoy (1998), "Distance Representations of Consumer Perceptions: Evaluating Appropriateness by Using Diagnostics," *Journal of Marketing Research*, 35 (May), 137–53.

Glazer, Rashi and Kent Nakamoto (1991), "Cognitive Geometry: An Analysis of Structure Underlying Representations of Similarity," *Marketing Science*, 10 (Summer), 205–228.

Gourville, John T. and Dilip Soman (1998), "Payment Depreciation: The Behavioral Effects of Temporally Separating Payments from Consumption," *Journal of Consumer Research*, 25 (2), 160–74.

Green, Paul E., J. Douglas Carroll, and S.M. Goldberg (1981), "A General Approach to Product Design Optimization via Conjoint Analysis," *Journal of Marketing*, 45 (3), 17–37.

Heath, Chip (1995), "Escalation and De-escalation of Commitment in Response to Sunk Costs: The Role of Budgeting in Mental Accounting," *Organizational Behavior and Human Decision Processes*, 62 (April), 38–54.

——— and Martin G. Fennema (1996), "Mental Depreciation and Marginal Decision Making," *Organizational Behavior and Human Decision Processes*, 62 (April), 38–54.

——— and Jack B. Soll (1996), "Mental Budgeting and Consumer Decisions," *Journal of Consumer Research*, 23 (June), 40–52.

Henderson, Pamela W. and Robert A. Peterson (1992), "Mental Accounting and Categorization," *Organizational Behavior and Human Decision Processes*, 51 (1), 92–117.

Hirst, D. Eric, Edward J. Joyce, and Michael S. Schadewald (1994), "Mental Accounting and Outcome Contiguity in Consumer-Borrowing Decisions," *Organizational Behavior and Human Decision Processes*, 58 (April), 136–52.

Johnson, Michael D. (1984), "Consumer Choice Strategies for Comparing Noncomparable Alternatives," *Journal of Consumer Research*, 11 (December), 741–53.

Joshi, Ashwin W. and Sanjay Sharma (2004), "Customer Knowledge Development: Antecedents and Impact on New Product Performance," *Journal of Marketing*, 68 (October), 47–59.

Kahn, Barbara E. and Robert J. Meyer (1991), "Consumer Multiattribute Judgments Under Attribute-Weight Uncertainty," *Journal of Consumer Research*, 17 (March), 508–522.

Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47 (March), 263–91.

Markman, Arthur B. and Douglas L. Gentner (1993), "Splitting the Differences: A Structural Alignment View of Similarity," *Journal of Memory and Language*, 32 (4), 517–35.

——— and Douglas L. Medin (1995), "Similarity and Alignment in Choice," *Organizational Behavior and Human Decision Processes*, 63 (2), 117–30.

Mukherjee, Ashesh and Wayne Hoyer (2001), "The Effect of Novel Attributes on Product Evaluation," *Journal of Consumer Research*, 28 (3), 462–72.

O'Curry, Suzanne and Michal Strahilevitz (2001), "Probability and Mode of Acquisition Effects on Choices Between Hedonic and Utilitarian Options," *Marketing Letters*, 12 (1), 37–49.

Okada, Erica Mina (2001), "Trade-Ins, Mental Accounting, and Product Replacement Decisions," *Journal of Consumer Research*, 27 (March), 433–46.

Peterson, Kim (2004), "The Name of the Game Is Software," *Seattle Times*, (February 16), D1.

Shocker, Alan D. and V. Srinivasan (1979), "Multiattribute Approaches to Product Evaluation and Generation: A Critical Review," *Journal of Marketing Research*, 16 (May), 159–80.

Simonson, Itamar, Ziv Carmon, and Suzanne O'Curry (1994), "Experimental Evidence on the Negative Effect of Product Features and Sales Promotions of Brand Choice," *Marketing Science*, 13 (1), 23–40.

Soman, Dilip and Amar Cheema (2001), "The Effect of Windfall Gains on the Sunk-Cost Effect," *Marketing Letters*, 12 (1), 51–62.

Thaler, Richard (1980), "Toward a Positive Theory of Consumer Choice," *Journal of Economic Behavior and Organization*, 1 (1), 36–90.

——— (1985), "Mental Accounting and Consumer Choice," *Marketing Science*, 4 (Summer), 199–214.

——— (1999), "Mental Accounting Matters," *Journal of Behavioral Decision Making*, 12 (3), 183–206.

Tversky, Amos (1977), "Features of Similarity," *Psychological Review*, 84 (4), 327–52.

Urban, Glen L. and John R. Hauser (1993), *Design and Marketing of New Products*, 2d ed. Upper Saddle River, NJ: Prentice Hall.

Zhang, Shi and Gavan J. Fitzsimons (1999), "Choice-Process Satisfaction: The Influence of Attribute Alignability and Option Limitation," *Organizational Behavior and Human Decision Processes*, 77 (3), 192–214.

——— and Arthur B. Markman (1998), "Overcoming the Early Entrant Advantage: The Role of Alignable and Nonalignable Differences," *Journal of Marketing Research*, 35 (November), 413–26.

Richard G. McFarland, Goutam N. Challagalla, & Tasadduq A. Shervani

Influence Tactics for Effective Adaptive Selling

The adaptive selling literature identifies effective salespeople as those who match their influence tactics to suit the characteristics of buyers. However, prior research is largely silent on the specific influence tactics that salespeople use and the effectiveness of these tactics across different types of buyers. The authors propose a theoretical model that uses Kelman's (1961) underlying influence processes of internalization, compliance, and identification to identify the seller influence tactics that salespeople use and to assess which of these tactics will resonate with three types of buyers: task-oriented buyers, interaction-oriented buyers, and self-oriented buyers. The authors test their model with data from 193 bidirectionally matched buyer–seller dyads. The results strongly support the theoretical model and suggest that buyers are more complex than originally presumed. However, salespeople seem to recognize this complexity and use the combination of influence tactics prescribed by theory for persuading these types of buyers.

Two widely accepted concepts in the sales literature are that (1) influence tactics are the mechanism through which salespeople persuade buyers in interactions (e.g., Brown 1990; Spiro and Perreault 1979) and (2) adapting sales tactics to different buyers is important for superior performance (e.g., Spiro and Weitz 1990; Sujan 1986; Weitz 1981). Notably, research at the intersection of these two notions is limited. As a consequence, little guidance is provided to salespeople regarding what influence tactics they should use with different types of buyers. This study sheds light on this important issue by proposing a theoretically grounded framework for investigating the effectiveness of sales influence tactics across groups of buyers who are dissimilar in their characteristics.

Frazier and Summers (1984) introduced the notion of influence strategies to the marketing literature. These influence strategies, or influence tactics, were developed for interorganizational, firm-to-firm channels contexts. Frazier and Summers (1984, p. 44) recognize the context for which their tactics were developed and specifically note that their tactics provide "researchers with an extended perspective on the influence process within channels of distribution. Of particular interest are boundary personnel, whose responsibilities lie in the areas of customer service and territorial management, as opposed to sales personnel." Whether Frazier and Summers's (1984) tactics are equally applicable in personal selling contexts remains in question. Thus, an objective of this study is to extend the work of these scholars and others (e.g., Boyle et al. 1992; Payan and McFarland 2005; Venkatesh, Kohli, and Zaltman 1995) and identify tactics that are especially pertinent in sales settings. We refer to these as "seller influence tactics" or, more simply, SITs.

Although identification of SITs is important, the adaptive selling literature suggests that it is insufficient unless it can be pinpointed which tactics should be used with different buyers. Sheth (1976) differentiates buyers on the basis of their "task orientation," "self orientation," and "interaction orientation."[1] Williams and Spiro (1985) validated this classification in sales settings; however, the influence tactics that will resonate with these three buyer orientations is unknown. Accordingly, a primary objective of the current study is to develop a theoretical framework for identifying the influence tactics that "work" with buyers of each of these three orientations. In addition to this theory development and its empirical testing, our study extends Sheth's work by identifying clusters of buyers on the basis of their orientations, and it examines the combinations of SITs that are used within each cluster.[2]

Our theoretical model argues that the three influence mechanisms that Kelman (1961) proposes—"internalization," "compliance," and "identification"—serve as the bridge between influence tactics and buyer orientations. We argue that influence tactics operate through one of these three underlying influence mechanisms. That is, some influence tactics operate through the mechanism of internalization, and others operate through compliance or identification. Furthermore, we argue that buyers with task, interaction, and self orientations respond differentially to

Richard G. McFarland is an assistant professor, Department of Marketing, Kansas State University (e-mail: mcfarlan@ksu.edu). Goutam N. Challagalla is an associate professor, College of Management, Georgia Institute of Technology (e-mail: gnc@gatech.edu). Tasadduq A. Shervani is an associate professor, Cox School of Business, Southern Methodist University (e-mail: tshervan@mail.cox.smu.edu).

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

---

[1] Sheth (1976) categorizes buyers in terms of task-oriented, interaction-oriented, and self-oriented communication styles. On the basis of the insightful comment of a reviewer, we refer to these communication styles as "orientations."

[2] We thank a reviewer for suggesting that we cluster buyers on the basis of their orientation scores and identify the combinations of influence tactics that are effective for each cluster.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

these three influence mechanisms. Thus, Kelman's underlying processes provide the key for unlocking the specific influence mechanisms that buyers will respond to and for identifying the influence tactics that operate through these mechanisms.

We use matched buyer–seller dyads to test the hypotheses. The use of dyadic data is consistent with the recommendations of several researchers, who state that influence tactics must be examined from the perspective of both parties, not just one (e.g., Achrol, Reve, and Stern 1983; Kim 2000). In addition, our study design reduces the chances of capitalizing on common method biases by having salespeople report their perceptions of a specific buyer's orientation and having that buyer report his or her perception of the SITs used by the same salesperson as well as the effectiveness of these SITs (Netemeyer et al. 1997; Ross, Anderson, and Weitz 1997).

We organize the article as follows: We first discuss the identification of the influence tactics in sales settings. Next, we link Kelman's (1961) underlying influence processes to these influence tactics. We then develop the hypotheses. We describe the data collection and analytic procedures, present our results, and then conclude.

Influence Tactics

Influence tactics have been examined at the interorganizational, firm-to-firm level of analysis in the marketing channels literature and at the intraorganizational, person-to-person level of analysis in the organizational behavior and buying-center literature. The personal selling context falls into a third category that can be described as an interorganizational, person-to-person relationship in which the unit of analysis is the individual relationship. Consequently, the influence strategies identified in the channels, organizational behavior, and buying-center literature need to be assessed for their applicability in sales settings.

Frazier and Summers (1984) define influence strategies as compliance-gaining tactics that channel members use to achieve desired actions from channel partners. Influence strategies "involve the alternative means of communication available to a firm's personnel in their influence attempts" (Frazier and Rody 1991, p. 52). Frazier and Summers (1984) propose six tactics: information exchange, recommendations, requests, threats, promises, and legalistic pleas (for brief definitions of the influence strategies from the marketing channels literature, see Table 1). These six tactics have been used extensively in the channels literature (e.g., Boyle and Dwyer 1995; Frazier and Summers 1986; Kim 2000). The same six tactics have also been examined in intrafirm contexts (e.g., Venkatesh, Kohli, and Zaltman 1995).

The literature identifies additional tactics to those that Frazier and Summers (1986) discuss that are likely to be especially pertinent in personal selling. These tactics focus on emotional utilities, which sales researchers have long acknowledged as having an important role in influencing buyers (e.g., Brown, Cron, and Slocum 1997; Dixon, Spiro, and Jamil 2001; Evans 1963; Webster 1968; Weitz 1981). For example, several studies emphasize the importance of ingratiation in sales contexts (e.g., Spiro and Perreault 1979; Strutton, Pelton, and Tanner 1996). Ingratiation focuses on influencing buyers by getting them to like the salesperson. Likewise, Yukl and Tracey (1992) note the importance of appealing to people's emotions and ideals through the use of inspirational appeals. Ingratiation and inspirational appeals are focused primarily on satisfying the psychological needs of the target. Salespeople's use of ingratiation and inspirational appeals is consistent with literature that notes the importance of attractiveness (e.g., Ahearne, Gruen, and Jarvis 1999), similarity (e.g., Crosby, Evans, and Cowles 1990), and liking (e.g., Doney and Cannon 1997) in professional selling.

TABLE 1
Marketing Channels Influence Strategy Definitions

Influence Strategy	Traditional Channels Definition
Information exchange	The source discusses general issues and procedures to try to alter the target's general perceptions without stating a request.
Recommendations	The source predicts that the target will be more profitable if the target follows the source's suggestions.
Requests	The source simply states the actions it would like the target to take.
Threats	The source threatens the target with a future penalty if the target does not comply with a request.
Promises	The source promises the target a reward if the target complies with a request.
Legalistic pleas	The source cites a legalistic, contractual, or informal agreement that requires the target to perform a certain action.

Notes: We drew these definitions from the work of Payan and McFarland (2005). We describe the definitions for the SITs in the text.

Identifying and Defining SITs

To pinpoint which of the eight influence tactics identified previously are appropriate in personal selling, we first conducted ten in-depth personal interviews with professional buyers and industrial salespeople. The purpose of these interviews was to allow participants to describe commonly used influence tactics, determine whether other influence tactics should be included in the SITs framework, help with item generation, and determine whether any of the scales we identified from prior research could be adapted to the personal selling context. In these interviews, we asked participants open-ended questions (e.g., "What type of tactics do professional salespeople use to influence buyers?") and specific questions about whether salespeople use a specific tactic, such as promises. We also sought feedback on measurement items from prior research.

We drew two conclusions from these interviews. First, four of the six influence tactics that Frazier and Summers (1984) define in the marketing channels literature formed a solid basis in the formation of SITs. Interviewees noted using information exchange, recommendations, threats, and promises; they regarded two tactics as either implicit (requests) or simply not used (legalistic pleas). Second, buyers and salespeople identified ingratiation and inspirational appeals as common in personal selling.

On the basis of these findings and further input from domain experts, we conducted a pretest with 17 executive MBA students involved in the purchasing function in their organizations. The purpose of this pretest was to confirm which tactics we should include and to assist with item purification. We provided participants with scales for the six influence tactics identified in the channels literature (including requests and legalistic pleas) and for ingratiation and inspirational appeals. Participants confirmed that the items for requests were confusing in the personal selling context because requests are implicit in sales interactions. They also confirmed that salespeople seldom, if ever, use legalistic pleas. Finally, participants confirmed that ingratiation and inspiration appeals are common sales tactics. Next, we define each of the six SITs.

Information exchange. Information exchange involves the communication of information, including asking questions, without making any specific recommendations to influence positively the buyer's general perceptions about the benefits of the salesperson's offerings. "The source firm's boundary personnel use discussions of general business issues … to alter the target's general perceptions;… however, no specific target action is requested" (Frazier and Summers 1984, p. 45).

Recommendations. Recommendations are arguments used to convince a customer that products or services purchased from the salesperson would be beneficial to the customer's firm (Fitzsimons and Lehmann 2004; Frazier and Summers 1984; Rackham 1988). Venkatesh, Kohli, and Zaltman (1995, p. 72) define recommendations as follows: "[T]he source [of influence] suggests that following a specific course of action is likely to be beneficial."

Threats. Threats are implied or stated negative sanctions that the salesperson asserts will be applied to the buyer's firm if the buyer does not comply with the seller's request (Boyle and Dwyer 1995; Boyle et al. 1992). "The source communicates to the target that it will apply negative sanctions should the target fail to" comply with the source's request (Frazier and Summers 1984, p. 46).

Promises. With promises, the seller "pledges to provide the target with a specific reward, contingent on the target's compliance with the source's" request (Frazier and Summers 1984, p. 46). Promises are pledges of future rewards for the buyer's firm. Venkatesh, Kohli, and Zaltman (1995, p. 73) state that "the deprivation of rewards may be considered to be equivalent to the imposition of sanctions." Thus, the influence tactic of promises is similar to that of threats, though it may be more tolerable than threats because the nonmediation of rewards leaves the buyer at the status quo, whereas the mediation of sanctions leaves the buyer worse off (Venkatesh, Kohli, and Zaltman 1995).

Ingratiation. Ingratiation involves the use of "behaviors that are designed to enhance one's interpersonal attractiveness" and improve rapport with the target of influence (Kumar and Beyerlein 1991, p. 619). These actions consist primarily of praising a customer for his or her achievements (other enhancement) and expressing attitude similarity (Kipnis and Schmidt 1988; Kipnis, Schmidt, and Wilkinson 1980).

Inspirational appeal. An inspirational appeal is "a request or proposal that arouses enthusiasm by appealing to a target's [higher-order] values, ideals, and aspirations" (Yukl and Tracey 1992, p. 640). Inspirational appeals focus on arousing a positive affective response in the target, which serves to motivate the target toward a desired response. Forgas (1995) argues that positive affective arousal leads to more positive expectancies and, thus, to higher levels of compliance. Inspirational appeals may attempt to motivate people to go beyond their personal interests for the greater good by appealing to higher-order psychological needs (Yukl and Tracey 1992).

Linking Influence Processes to SITs

Kelman (1958, 1961) and others (e.g., Bearden, Netemeyer, and Teel 1989) describe three underlying processes of interpersonal influence: internalization, compliance, and identification. Each of the six SITs operates through one of these mechanisms of interpersonal influence. Specifically, we propose that information exchange and recommendations operate through the process of internalization, threats and promises operate through the mechanism of compliance, and ingratiation and inspirational appeals operate through the processes of identification.

Information exchange and recommendations tactics attempt to alter a buyer's "perceptions regarding the inherent desirability of the intended behavior" (Frazier and Summers 1984, p. 44). We propose that these tactics operate through the influence process of internalization, which occurs when the target of influence (i.e., the buyer) does what the source of influence (i.e., the salesperson) requests because the target believes that the action is appropriate and is in his or her own best interest (Kelman 1961). A target accepts the induced behavior because it is in agreement with his or her value system or because the content of the behavior is intrinsically rewarding (Smith and Barclay 1997). In sales settings, internalization occurs when buyers accept influence because they find the content of the influence to represent a useful solution or because it is congruent with their own orientations (e.g., approach to problem solving).

In contrast, threats and promises tactics focus on changing the motivation of the customer to comply on the basis of either positive or negative source-controlled consequences (Frazier and Summers 1984; Keith, Jackson, and Crosby 1990; Venkatesh, Kohli, and Zaltman 1995). We propose that these tactics operate through the process of compliance, which occurs when people accept influence (or change their beliefs) because they hope to obtain a favorable reaction or reward from another person (Kelman 1961). The recipient may be interested in attaining certain rewards or in avoiding certain "punishments" that the influencing agent controls (Brown, Lusch, and Nicholson 1995; Gundlach and Cadotte 1994). For example, buyers may be receptive to "compliance" because they hope to obtain preferential delivery of a product or desirable credit terms. The recipient of the influence likely does not adopt the induced behavior because he or she believes in its content but rather because it is instrumental in goal accomplishment.

Ingratiation and inspirational appeals aim to satisfy a target's psychological needs by being attractive, by maintaining a fulfilling relationship with the target, and through eliciting positive emotional responses. We propose that these tactics operate through the process of identification, which occurs when a target is receptive to influence because he or she finds it attractive to maintain a satisfying or self-defining relationship with the influencer (Kelman 1961). Identification is based on satisfying psychological needs. A person can try to establish or maintain a relationship with another party through identification in different ways. For example, he or she can "identify" with another by echoing what the other person expresses, endorsing that person's views, or even behaving like that person, thus forming a satisfying or self-defining relationship with that person. In summary, ingratiation and inspirational appeals tactics operate through the process of identification because they seek to elicit enthusiasm, liking, and other positive emotive feelings toward the salesperson, the salesperson's firm, and/or his or her offerings.

Hypotheses

As we noted previously, Sheth (1976) suggests that buyers can be distinguished on the basis of their task orientation, self orientation, and interaction orientation. These buyer orientations are rooted in the literature on leadership (Bass 1960). Sales researchers examining customers' social styles propose similar conceptualizations (Comstock and Higgins 1997; Williams and Spiro 1985). For example, Blake and Mouton (1980) propose that customers may have two orientations: concern with the sale and concern about the other person. These two orientations parallel task and interaction buyer orientations. Furthermore, customers who are low on both of these orientations may be categorized as having a self orientation (Williams and Spiro 1985).

Buyer orientations have also been referred to as communication styles in the literature because they represent a person's overall communication pattern and goal expectations from social interactions (Williams and Spiro 1985). Sheth (1976, p. 382) defines these orientations as "representing the format, ritual or mannerism which a buyer" adopts in his or her interactions with salespeople. A buyer orientation is rooted in personality variables, socialization processes, personal lifestyles, and situational factors (Sheth 1976), suggesting that though a buyer's orientation is not a personality trait, it can be relatively stable. This has two important implications. First, because buyer orientations are somewhat stable, it is possible to identify a buyer's orientation and communication preferences. Second, the adaptive selling literature suggests that sellers who use the influence tactics that are most appropriate for a particular buyer orientation are likely to reap greater benefits than those who do not.

SITs' Effectiveness and Buyer Task Orientation

Buyers with a task orientation are goal oriented and purposeful (Sheth 1976). They want to accomplish the task at hand as efficiently as possible, and they are intolerant of any activity that is inefficient and non–goal oriented or that deviates from the current task on which they are focusing (Sheth 1976). Williams and Spiro (1985) suggest that people with a task orientation tend to be mechanistic in their approach to others and wish to make the best decision possible for their organization by being focused on the task. People with a task orientation are more focused on the task and less focused on relationship building than people with an interaction orientation. Similarly, those with a task orientation are more focused on achieving organizational objectives than those with a self orientation (Miles, Arnold, and Nash 1990).

Salespeople who rely on SITs that operate through the process of internalization are likely to achieve greater manifest influence with buyers with a high task orientation. This is because such buyers need to be convinced of the merit of an idea or action and observe how it meets organizational objectives. They need to believe that the solutions being proposed are appropriate and congruent with their value system. Information exchange and recommendations, which operate through the process of internalization, are likely to appeal to buyers with a high task orientation because these tactics attempt to convince the target of the inherent attractiveness of the intended action by presenting facts and logical arguments (Frazier and Summers 1984). Indeed, prior research suggests that buyers with a task orientation respond more favorably to influence attempts that are high in knowledge and information than to attempts that are low in knowledge and information (Williams and Spiro 1985).

More specifically, the higher the information exchange, the more it educates buyers about the salesperson's products and services, general business issues, and operating procedures. Furthermore, when information exchange is high, it enhances buyers' knowledge and enables them to evaluate different solutions. Buyers with a high task orientation will value such information because it helps them achieve organizational goals. Similarly, recommendations are likely to appeal to such buyers. Influence tactics laden with recommendations provide specific suggestions and arguments to the buyer for making purchase decisions (Frazier and Summers 1984). They even suggest the behaviors that the target should follow to accomplish organizational objectives. Thus, recommendations help buyers with a high task orientation compare and contrast different solutions efficiently, enabling them to make decisions more expeditiously.

> H_1: The higher a buyer's task orientation, the stronger is the positive effect of salesperson (a) information exchange and (b) recommendations on buyer manifest influence.

SITs' Effectiveness and Buyer Self Orientation

Buyers with a self orientation are concerned almost entirely with their own welfare. For them, the sales interaction is "merely the theater in which ... needs can be satisfied" (Bass 1960, p. 149). They view the interaction from a selfish point of view (Sheth 1976). They are preoccupied with themselves during interactions with others and are less empathetic toward other people (Miles, Arnold, and Nash 1990). Success in meeting personal needs outweighs the needs for task effectiveness and for accomplishing organizational objectives. Neither interaction nor task effectiveness is rewarding to the self-oriented individual, unless such effectiveness contributes to his or her self-esteem, status, or other personal gains (Bass 1960).

Because buyers with a self orientation are keen on promoting their own agenda and interests, they are likely to be most responsive to the influence process of compliance. Influence processes based on compliance aim to obtain a change in behavior not by altering perceptions about the desirability of the intended action but through the use of rewards or sanctions. Such processes are effective in obtaining behavioral change when targets are either reluctant to act or will not act unless they personally benefit from taking an action (Kelman 1961). Threats and promises, which operate through the process of compliance, are likely to be most effective with buyers with a high self orientation because they provide inducements for a necessary behavior.

When buyers go too far in preserving their own self-interest to the detriment of the buyer–seller relationship or to the detriment of the buyer's own organization, the use of threats (e.g., withdrawing preferential delivery dates or applying stringent credit terms) may be the only effective means to motivate them to act. In addition, because buyers with a high self orientation are primarily interested in their own benefits and are highly concerned with extrinsic rewards (Williams and Spiro 1985), promises are likely to appeal to these buyers because it rewards them with incentives and concessions (Kelman 1958). In addition, because such buyers often display a reluctance to act (Bass 1960), promises and threats are a powerful means for overcoming this reluctance. Collectively, these arguments indicate that buyers with a self orientation are likely to be responsive to the influence tactics of threats and promises.

We also anticipate that ingratiation influence tactics will be effective. Because buyers with a self orientation are self-absorbed, they like to receive attention and brag about their accomplishments. At the same time, they are motivated in interactions to gain esteem (Stotland and Zander 1958). They tend to display dependent and succorant behavior; in other words, they need validation and support from others (Bass 1967). A key aspect of ingratiation tactics is other enhancements. Ingratiation serves the dual role of providing buyers the validation they seek and supporting their belief of self-importance.

> H_2: The higher a buyer's self orientation, the stronger is the positive effect of salesperson (a) threats, (b) promises, and (c) ingratiation on buyer manifest influence.

SITs' Effectiveness and Buyer Interaction Orientation

Buyers with an interaction orientation believe that socializing is a critical aspect of the interaction process. They are interested in forming friendships and fostering interpersonal relationships (Bass 1960). Such buyers put more emphasis on relational elements than on the specifics of the transaction. They are "often compulsive in establishing a personal relationship" with the salesperson and only then "[get] involved in the specific content of interaction[s]" (Sheth 1976, p. 385). Developing personally satisfying relationships is paramount to such buyers. Given the focus on social interactions and relationships, buyers with an interaction orientation are likely to be most responsive to the influence process of identification. Indeed, Kelman's (1961) theory suggests that people are more likely to be influenced through the identification process when they are concerned with social anchorage.

Ingratiation and inspirational appeals, which rely on identification as an influence process, are likely to result in greater manifest influence among buyers with a high interaction orientation. Importantly, ingratiation influence tactics are aimed at building a rapport with buyers (Kumar and Beyerlein 1991). Salespeople who rely on ingratiation tactics behave affably, discuss shared interests and hobbies, and give praise and compliments (Kipnis and Schmidt 1988). These types of behaviors help forge interpersonal relationships, which are consistent with the goals of a buyer with an interaction orientation. Conversely, salespeople who hardly rely on ingratiation tend to focus less on building relationships or on appealing to the buyer's positive image. These salespeople are likely to have lower levels of manifest influence with such buyers.

Buyers with an interaction orientation prefer sales tactics that emphasize emotional utilities, which indicates that such buyers look beyond product features and product benefits and make decisions based on emotions (Williams, Spiro, and Fine 1990). This concept is consistent with recent consumer research that indicates that some people use a process in decision making known as the "how-do-I-feel-about-it" heuristic (Adaval 2001; Pham et al. 2001). Inspirational appeals rely on attaching emotional relevance to a product, independent of the product's performance or functional utility. Finally, inspirational appeals are likely to resonate with buyers with an interaction orientation because such appeals seek to elicit excitement and other positive emotions in regard to the salesperson and/or his or her offerings.

> H_3: The higher a buyer's interaction orientation, the stronger is the positive effect of salesperson (a) ingratiation and (b) inspirational appeals on buyer manifest influence.

SITs Usage Based on Buyer Orientations

Prior studies indicate that, in general, salespeople are aware of the type of sales interactions buyers prefer (Comstock and Higgins 1997) and use this knowledge when interacting with buyers (Spitzberg and Cupach 1984). Research suggests that the choice of influence tactics is likely to depend on the characteristics of the target of influence (Erez, Rim, and Keider 1986). In support of this line of reasoning, empirical evidence suggests that people use influence tactics that they believe are most effective in gaining compliance from others (Leigh and McGraw 1989; Venkatesh, Kohli, and Zaltman 1985). Collectively, these arguments suggest that salespeople will vary the frequency with which they use SITs in an attempt to conform to buyers' preferred orientations.

Consistent with the preceding logic, because buyers with an interaction orientation are likely to respond to influence attempts that rely on internalization, it follows that salespeople will use information exchange and recommendations with such buyers. Similarly, when dealing with buyers with a self orientation, salespeople will use influence attempts that rely on obtaining compliance. That is, they will use threats and promises. In addition, they will use ingratiation because such buyers also respond to influence attempts that enhance their self-image. Finally, salespeople will use ingratiation and inspirational appeals when interfacing with buyers who are high on interaction orientation because attempts that rely on identification as an influence mechanism resonate with such buyers.

> H_4: The higher the task orientation of a buyer, the greater is a salesperson's use of (a) information exchange and (b) recommendations; the higher the self orientation of a buyer, the greater is a salesperson's use of (c) threats, (d) promises, and (e) ingratiation; and the higher the interaction orientation of a buyer, the greater is the salesperson's use of (f) ingratiation and (g) inspirational appeals.

Method

Common Method Biases and Dyadic Study Design

Our methodology and sample design are consistent with Podsakoff and colleagues' (2003) recommendations for reducing common methods biases. They recommend (1) obtaining measures of the predictor and criterion variables from difference sources; (2) offering anonymity/confidentiality to reduce the chances of responses that are socially desirable, lenient, acquiescent, and consistent with how participants think researchers want them to respond; and (3) informing the respondents that there are no correct or incorrect answers and that they should respond as honestly as possible to reduce evaluation apprehension. In our study, we used bidirectionally matched dyads such that the key variables are non-self-reported. That is, we had buyers report on the influence tactics that salespeople use, and we had salespeople report on buyer orientations. In addition, we informed the respondents that their responses were confidential and that there were no right or wrong answers to the items in the survey.

Final Pretest of SITs Measures

We conducted a final pretest on the SITs measures using a convenience sample of 93 salespeople working for a *Fortune* 500 company that manufactures farm equipment. Before administering the pretest questionnaire, we asked two executives from the participating corporation to review the questionnaire and confirm that all instructions and questions were clear and unambiguous. We conducted exploratory factor analysis on the data to investigate the unidimensionality and factor structure of the sales tactics constructs (e.g., Challagalla and Shervani 1996; Singh and Rhoads 1991). Using a cutoff eigenvalue of 1.0, we found a six-factor solution, consistent with our expectations. The reliability of each measure exceeded the .70 threshold for acceptable reliability, which suggests that the measures were internally consistent (Nunnally 1978). At this stage, we determined that the measures were ready for the final study.

Dyadic Sample and Data Collection Procedure

The unit of analysis in this study is the individual dyadic relationship between buyers and sellers. We chose an industrial retail setting in which a dealer salesperson sells to end users. The dealers in the study were resellers of agricultural equipment for a *Fortune* 500 company. These dealers represent the manufacturer's products exclusively. The dealers operate in a competitive environment in which customers can choose from multiple brands. Most customers are small-business operators. The customer's purchase decision is most often made by an individual rather than by a committee.

We first approached the vice president of dealer operations at the agricultural equipment manufacturer. This company divided its market into 100 regions and had multiple dealers in each region. Our sample frame consisted of 400 salespeople, one from each of 400 dealerships (4 dealerships per region). Thus, the maximum number of independent, bidirectionally matched dyads that could be achieved was 400. Following the procedure we describe subsequently, we collected 193 independent, bidirectionally matched, usable dyads, thus yielding a 48% achieved dyadic response rate.

We used a two-phase process to obtain this dyadic data. In the first phase, we obtained a customer list for each dealership from the dealers' parent organization. This list included the names of three customers per dealership. These customers were chosen at random from each dealership's customer database. We eliminated 12 because of inaccurate addresses, missing or incorrect customer names, and duplicate entries. To ensure accurate responses, we guaranteed respondents confidentiality and informed them that there were no correct or incorrect responses. We asked customers to write the name of the salesperson they had primary contact with at the dealership on their questionnaire and to respond to all questions with that salesperson in mind. After three waves of mailings, we received 459 responses, for a response rate of 39%. In some cases, we received multiple responses for each dealership. The 459 customers who responded to the survey mentioned a total of 290 unique dealer salespeople.

In the second phase, we mailed questionnaires to the 290 salespeople that customers in the first phase identified. To ensure that we obtained matched dyads, we provided each salesperson with the name of the customer who identified him or her and asked the salesperson to respond to all questions with that specific customer in mind (we did not reveal customer responses). When two or more customers identified the same salesperson as their primary contact, we selected one customer response on the basis of an assigned number chosen at random, using a random-number generator. This ensured that each salesperson responded to just one questionnaire in which a specific customer was identified. This also ensured that each dyad was independent in the final analysis.

We eliminated seven salespeople from the sample frame because of incorrect addresses and turnover. This left an effective sample of 283 salespeople. We received 199 responses, for a 70% response rate within this sample. Of these, we eliminated 6 responses because of missing data, for a total of 193 full-data salesperson responses and, consequently, 193 individually matched independent pairs of buyer–seller dyads, each from a different dealership. We tested nonresponse bias with the procedure that Armstrong and Overton (1977) outline, using a median split between early and late responders for each sample. There were no significant differences between early and late responders in either sample.

Measures

The items and scale anchors for the measures appear in the Appendix. The summary statistics for all variables used in hypotheses testing appear in Table 2. Because of the likely response bias in salespeople's self-reported influence tactics, buyers responded to the influence tactics items. Following the procedures we discussed in detail previously, we modified Schriesheim and Hinkin's (1990), Venkatesh, Kohli, and Zaltman's (1995), and Yukl and Tracey's (1992) scales for context and unit of analysis. We used Kohli and Zaltman's (1988) manifest influence scale to measure the effectiveness of each influence tactic. Again, we modified the items for context. To eliminate the possibility of social desirability bias, buyers responded on this scale.

We used William and Spiro's (1985) multi-item scales for task orientation, self orientation, and interaction orientation. Salespeople were the respondents for these measures because their perceptions of the buyer's orientation drive their selection and use of SITs.

Analytic Procedures and Results

Construct Validity

We analyzed the SITs, buyer orientations, and manifest influence measures simultaneously in a single confirmatory factor analysis using LISREL 8.54 (Jöreskog and Sörbom 1996). Fit exceeded the standard cutoff criteria: $\chi^2 = 505$, d.f. = 419; root mean square error of approximation = .033; comparative fit index = .99; and standardized root mean square residual = .045. The composite trait reliabilities of each factor exceeded the recommended cutoff criteria of .70 (except for buyer interaction orientation, which was .68), and the variance extracted for each factor exceeded the .50 cutoff criteria (Fornell and Larcker 1981; Hair et al. 1998). Together, this provides strong evidence that the measures are internally consistent. All indicators loaded significantly on their corresponding latent factor, indicating convergent validity. Using the procedure that Fornell and Larcker (1981) recommend, we obtained discriminant validity for all pairs of measures.

Testing the Effectiveness of SITs

H_1–H_3 predicted that the effectiveness of each SIT would vary with each buyer orientation. To test these moderator effects, we split the sample along the median of the moderator for each of the three orientation variables to obtain "high" and "low" subgroups (Johnston 1984; Kohli 1989). For each of the three orientations, we then regressed manifest influence on the six SITs using cases from both subgroups (restricted run). Next, we ran a second regression in which we allowed the regression coefficients of the six influence tactics to take different values in the high and low subgroups. By including all six influence tactics, we were able to examine not only whether the hypothesized relationships were in the expected direction but also whether the nonhypothesized relationships were nonsignificant.

We performed a Chow test (see Johnston 1984) to assess whether the difference in the sum of squared residuals in the restricted and unrestricted run was significant. The Chow test results indicate that task orientation ($F_{(6, 189)}$ =

TABLE 2
Correlation Matrix and Summary Statistics

Measure	1	2	3	4	5	6	7	8	9	10
1. Buyer task orientation	(.87)									
2. Buyer interaction orientation	.31	(.67)								
3. Buyer self orientation	–.39	–.20	(.89)							
4. Information exchange	.13	.17	–.15	(.87)						
5. Recommendations	.04	.15	–.07	.70	(.90)					
6. Threats	–.26	–.24	.20	–.21	–.16	(.96)				
7. Promises	–.32	–.19	.27	.02	.12	.40	(.82)			
8. Ingratiation	–.04	.18	.05	.53	.51	–.06	.22	(.78)		
9. Inspirational appeals	.03	.26	–.05	.49	.58	–.13	.16	.66	(.80)	
10. Manifest influence	–.10	.10	.00	.54	.53	–.04	.20	.47	.42	(.92)
M	4.00	3.52	1.90	3.74	3.31	1.18	1.55	3.31	3.56	3.40
SD	.72	.90	.88	1.05	1.19	.59	.81	1.03	1.04	.98
Composite reliability	.87	.68	.89	.87	.90	.97	.83	.77	.81	.92
Variance extracted	.68	.52	.72	.69	.75	.88	.62	.53	.58	.70

Notes: Cronbach's alphas appear on the diagonal of the correlation matrix. Correlations ≥ ±.14 are significant at $p < .05$

2.93, $p < .01$; adjusted $R^2 = .44$), self orientation ($F_{(6,\ 189)} = 2.20$, $p < .05$; adjusted $R^2 = .47$), and interaction orientation ($F_{(6,\ 188)} = 2.68$, $p < .01$; adjusted $R^2 = .46$) moderate the relationship between the six influence tactics and manifest influence. This finding lends overall support to the notion that buyer orientations matter when influence tactics usage is considered. The significant standardized regression coefficients from these analyses appear in Table 3.

As H_{1b} predicted, recommendations have a positive relationship to manifest influence when buyer task orientation is high ($b = .54$, $p < .01$), but recommendations are unrelated to manifest influence when buyers have a low task orientation. Contrary to H_{1a}, information exchange is positively related to manifest influence both when buyer task orientation is low ($b = .78$, $p < .001$) and when it is high ($b = .44$, $p < .05$), suggesting a main effect for this tactic. Thus, H_1 receives partial support. Notably, promises, ingratiation, and inspirational appeals are unrelated to manifest influence when buyer task orientation is high, whereas threats have a negative effect ($b = -.39$, $p < .05$). Overall, these findings indicate reasonable support for our theoretical model.

As H_{2b} and H_{2c} predicted, promises ($b = .21$, $p < .05$) and ingratiation ($b = .74$, $p < .001$) are positively related to manifest influence for buyers with a high self orientation, but they are unrelated to it for buyers with a low self orientation. Threats are unrelated to manifest influence for high–self orientation buyers, as H_{2a} predicted, but threats are negatively related to it for low–self orientation buyers ($b = -.30$, $p < .05$). The other influence tactics, except for recommendations, are unrelated to manifest influence when buyer self orientation is high. In general, these findings support H_2 and are consistent with our theoretical model.

Finally, H_3 receives strong support; both ingratiation ($b = .46$, $p < .01$) and inspirational appeals ($b = .39$, $p < .01$) are positively related to manifest influence when buyer interaction orientation is high. Notably, inspirational appeals are negatively related to manifest influence ($b = -.43$, $p < .05$) for buyers who are low in interaction orientation. We also find strong support for our theoretical model because none of the other influence tactics are related to manifest influence when buyer interaction orientation is high.

Buyer Orientation Clusters and SITs Usage

We used a two-stage analytical procedure to test H_4. In the first step, rather than assume that buyers exhibit only one orientation, we clustered the buyers using their scores on the three buyer orientations as input.[3] We employed hierarchical clustering using Wards method to cluster the buyers initially. We uncovered a three-cluster solution using the procedures that Milligan and Cooper (1985) recommend for determining the number of clusters. Next, as Punj and Stewart (1983) recommend, we used the cluster seeds from the hierarchical cluster analysis as starting values in a k-means cluster analysis. We eliminated four outliers on the basis of their distance from the cluster seeds.

Figure 1 depicts the plot of the cluster means. The results of the cluster analysis are interesting and suggest that buyers are more complex than originally presumed.

---

[3] Again, we thank a knowledgeable reviewer for recommending that we cluster the buyers on the basis of their orientations.

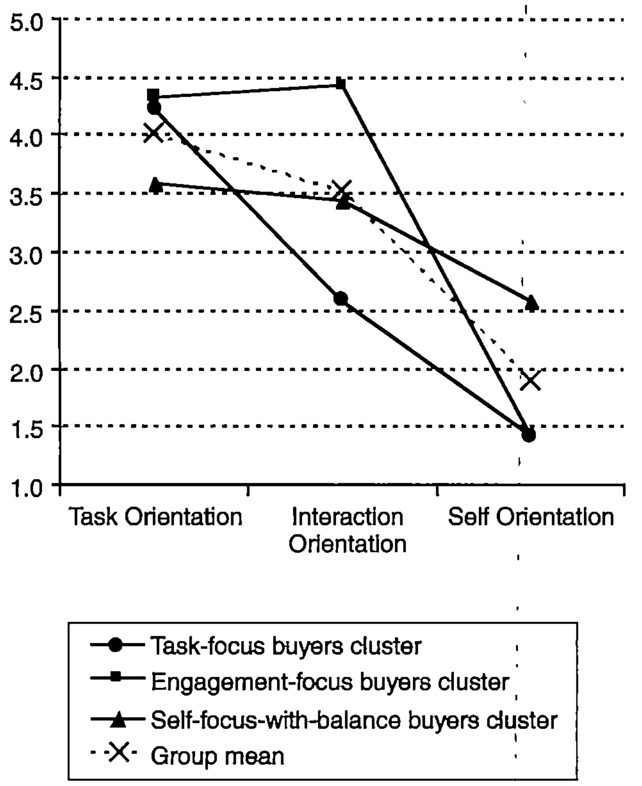
TABLE 3
Regression Coefficients Across High and Low Levels of the Buyer Orientations (Dependent Variable: Manifest Influence)

			Moderator Level			
Moderator Variable	**Independent Variables**	**Hypothesis**	**High**	**Low**	**Adjusted R^2**	**Chow Test**
Buyer Task Orientation	Information exchange	H_{1a}	.44*	.78***	.44	$F_{(6,\ 189)} = 2.93$**
	Recommendations	H_{1b}	.54**	—		
	Threats		–.39*	—		
	Promises		—	.17*		
	Ingratiation		—	.34*		
	Inspirational appeals		—	—		
Buyer Self Orientation	Information exchange		—	.75***	.47	$F_{(6,\ 189)} = 2.20$*
	Recommendations		.51**	.50**		
	Threats	H_{2a}	—	–.30*		
	Promises	H_{2b}	.21*	—		
	Ingratiation	H_{2c}	.74***	—		
	Inspirational appeals		—	—		
Buyer Interaction Orientation	Information exchange		—	1.05***	.46	$F_{(6,\ 188)} = 2.68$*
	Recommendations		—	—		
	Threats		—	—		
	Promises		—	—		
	Ingratiation	H_{3a}	.46**	—		
	Inspirational appeals	H_{3b}	.39*	–.43*		

*$p < .05$.
**$p < .01$.
***$p < .001$.
Notes: Dash indicates coefficient is not significant at the .05 level

Only one of the clusters aligns on a single orientation; the remaining clusters align on multiple orientations. About one-third of the buyers fell into each of the clusters, which we named "task-focus buyers," "engagement-focus buyers,"[4] and "self-focus-with-balance buyers" (see Table 4). Task-focus buyers have a much higher task orientation score (4.2) than interaction (2.7) and self (1.4) orientation scores. Engagement-focus buyers are high on both task (4.3) and interaction (4.4) orientations but are low on self orientation (1.4). Self-focus-with-balance buyers have higher scores on self orientation (2.6) than the other two clusters (i.e., task-focus buyers = 1.4 and engagement-focus buyers = 1.4). However, no orientation seems to be high within this third cluster (i.e., task orientation = 3.6, interaction orientation = 3.4, and self orientation = 2.6).

FIGURE 1
Plot of Mean Scores for Buyer Orientation Clusters



In the second step of the analysis, we used the three buyer clusters as the multicategory variable in a multivariate analysis of variance (MANOVA); the six SITs served as the independent variables. Table 4 depicts the mean scores of the SITs from the MANOVA. Wilks' lambda is .82 (F = 3.13, $p < .001$), Hotelling's T is .21 (F = 3.17, $p < .001$), and Roy's largest root is .17 (F = 5.20, $p < .001$) for the MANOVA, indicating that the mean values for the SITs vary significantly across groups. As Table 4 depicts, the univariate one-way analysis of variance (ANOVA) results indicate that the means for all SITs, except for that of recommendations, are significantly different across the clusters.

Helmert contrasts indicate that information exchange is in the significantly higher range for both the task-focus and the engagement-focus buyers and is in the lower range for the self-focus-with-balance buyers, which is consistent with H_{4a}. There are no differences in the use of recommendations across the clusters, which is not supportive of H_{4b}. The use of threats and promises is significantly higher for the self-focus-with-balance cluster than for the other two clusters, which is consistent with H_{4c} and H_{4d}. The use of ingratiation and inspirational appeals is significantly higher for the engagement-focus cluster than for the other two clusters, which is consistent with H_{4f} and H_{4g} but is inconsistent with H_{4e}. Overall, these findings indicate that salespeople use the influence tactics recommended by theory in their interactions with buyers.

Table 4 also points to the specific combination of SITs used in each cluster. For example, when dealing with engagement-focus buyers, salespeople use information exchange, recommendations, ingratiation, and inspirational appeals at a higher level, and the use of the latter two tactics is significantly higher than in the other clusters. Likewise, with self-focus-with-balance buyers, it might be expected

TABLE 4
Means for Buyer Orientation Clusters (MANOVA)

	Buyer Orientation Clusters[a]			ANOVA Results[b]	
	Task-Focus Buyers	Engagement-Focus Buyers	Self-Focus-with-Balance Buyers	F Statistic	*p* Value
Information exchange	**3.8**	**4.0**	3.5	3.40	< .05
Recommendations	3.4	3.5	3.2	.89	n.s.
Threats	1.2	1.0	**1.3**	4.32	< .05
Promises	1.5	1.2	**1.8**	7.44	< .001
Ingratiation	3.1	**3.6**	3.2	3.28	< .05
Inspirational appeals	3.4	**3.8**	3.4	3.42	< .05
Number of Observations	59	63	67		

[a] Mean values in bold are significantly higher than the other mean values along the same row, based on the Helmert comparison procedure ($p < .05$).
[b] These are the results from a one-way ANOVA with the clusters as the grouping variable.
Notes: The abbreviation "n.s." indicates that the ANOVA is not significant at the 05 level

---

[4] We thank one of the reviewers for alluding to this term.

that *all* influence tactics are used at moderate levels. This is indeed the case, except for threats and promises, which do not seem to be used much across the board. Nevertheless, the use of threats and promises in this cluster is higher than in the other clusters. Thus, overall, salespeople appear to use the combinations of tactics that would be expected on the basis of theory. The overall pattern of results from the MANOVA leads to two important conclusions: First, salespeople use combinations of tactics, and second, in general, the tactics they use follow the theoretical prescriptions of our model.

Post Hoc Analysis

To validate that the SITs effectiveness results (i.e., H_1–H_3) we obtained previously for each buyer orientation factor also hold true for the buyer clusters, we conducted a multigroup analysis using LISREL 8.54. Table 5 depicts these results. To assess whether the differences among groups are significant, we compared this model with a null model in which the gamma paths were fixed across each group. The chi-square difference ($\Delta\chi^2 = 55.89$, d.f. = 12) between these two models is significant at $p < .0001$, indicating that the null model should be rejected and that group differences are significant.

As we discussed previously, we find general support for H_1–H_3, which suggests that matching sales tactics to buyer orientations positively affects manifest influence. We find additional support for this notion, even when we examine buyer clusters (see Table 5). Indeed, seven of the nine paths are significant and in the expected direction. For example, consistent with H_1, we find that information exchange and recommendations resonate with clusters that include buyers with a task orientation (i.e., task-focus buyers and engagement-focus buyers). The two exceptions were inspirational appeals and threats, which had no impact on manifest influence for engagement-focus buyers and self-focus-with-balance buyers, respectively. In addition, unexpectedly, information exchange had an impact on the manifest influence of self-focus-with-balance buyers. Overall, for the hypothesized paths, with the exception of inspirational appeals, the results from the buyer clusters analysis are consistent with the Chow test results we present in Table 3.

Discussion

This study advances our knowledge of the use of influence tactics in buyer–seller interactions. It identifies the specific influence tactics that salespeople use with buyers and the relative effectiveness of these tactics across different buyers. Importantly, it demonstrates that salespeople match their use of SITs to different buyer orientations in a theoretically prescribed manner and, by doing so, that salespeople are able to influence positively the buyers' decision making.

Contributions to Theory

The importance of adaptive selling is widely recognized in the sales literature. This literature suggests that rather than use the same tactics, salespeople should adapt their tactics and that these adaptations should "suit" the buyers with whom they are dealing (Szymanski 1988; Weitz, Sujan, and Sujan 1986). After all, merely changing tactics is insufficient if the adaptations do not appeal to buyers. Notably, prior research is largely focused on the degree to which salespeople change their tactics and not on whether the adaptations are suitable. Our study makes a contribution to the literature by developing a theoretically grounded framework for understanding the suitability of each SIT for different types of buyers. Our theoretical model uses Kelman's (1961) influence processes to identify the influence tactics that will resonate with buyers on the basis of their orientations. By using theory to pinpoint the appropriate tactics for different buyers, we add to the stream of research on adaptive selling.

This study also contributes to the influence strategy literature by identifying the influence strategies, or tactics, that are used in sales contexts. Frazier and Summers (1984) developed what many researchers consider the most widely used taxonomy of influence strategies in marketing. However, they developed their strategies for a different context and for a different level of analysis; principally, these strategies are aimed at achieving influence within existing firm-to-firm-level relationships. In contrast, the personal selling context often deals with individual-level relationships. In this context, we find that it is necessary to add two tactics to Frazier and Summers's taxonomy, namely, ingratiation and

TABLE 5
Multigroup LISREL: The Impact of SITs on Manifest Influence by Buyer Cluster (Standardized Gamma Path Coefficients)

		Buyer Orientation Clusters		
Path from	**Path to**	**Task-Focus Buyers**	**Engagement-Focus Buyers**	**Self-Focus-with-Balance Buyers**
Information exchange	Manifest Influence	.31**	.28**	.45*
Recommendations	Manifest Influence	.25*	.21*	.11
Threats	Manifest Influence	–.02	.09	.03
Promises	Manifest influence	.10	.09	.23**
Ingratiation	Manifest influence	.17	.32*	.37*
Inspirational appeals	Manifest Influence	–.11	.03	–.12

*$p < .01$.
**$p < .001$.
Notes: $\chi^2_{(763)} = 2921.43$ for base model.

inspirational appeals. In addition, we find that Frazier and Summers's requests and legalistic pleas strategies are either implicit or uncommon in the interpersonal sales context. This updated taxonomy provides sales researchers with an appropriate context-specific set of influence tactics (i.e., SITs) for examining the salesperson influence process.

This study extends the work on buyer orientations (e.g., Sheth 1976; Williams and Spiro 1985) and helps identify clusters of buyers as they exist in industrial buying. Two interesting conclusions can be drawn from our study on this topic. First, buyers are more complex than previously believed. Our study uncovered three clusters of buyers. Although a dominant task orientation defines the first cluster of buyers, the other two clusters are not as straightforward. The second cluster of buyers, engagement-focus buyers, are high on both task *and* interaction orientation. Intriguingly, the third cluster of buyers, self-focus-with-balance buyers, is defined by moderate levels of all three orientations but with higher self orientation scores than the other two clusters. Second, the theory at the factor level (i.e., orientation level) appears to be robust. This means that salespeople can influence buyers as long as they use the tactics that theory recommends for each of the orientations a buyer reflects. Overall, our study deepens our understanding of how buyer orientations might cluster and the combinations of tactics to use with these different clusters of buyers.

Managerial Implications

Mangers are interested in knowing whether there are universally effective tactics and whether there are any tactics that should be avoided. This study provides interesting insights into this question by integrating the findings of several different analyses and perspectives on buyer orientations. If the traditional view is taken that buyers have a single dominant orientation, it would appear that *no* tactic is universally effective across all buyers. Solely on the basis of traditional analytic results, the following advice would be given: For buyers with a task orientation, sellers should use information exchange and recommendations; for buyers with an interaction orientation, sellers should rely on ingratiation and inspirational appeals; and for buyers with a self orientation, sellers should depend on promises and ingratiation. These results strongly support the key premise of the adaptive selling literature (e.g., Weitz, Sujan, and Sujan 1986) and suggest that a buyer's orientation should drive the use of influence tactics.

However, our additional analyses suggest that, more often than not, buyers align along multiple orientations. Indeed, only one-third of the buyers align along a single dominant orientation (i.e., task orientation); the other two-thirds have multiple orientations. This implies that if an orientation is ubiquitous among buyers, it is possible to uncover universally effective tactics. Notably, we find that buyers in all three clusters have moderate to high levels of task orientation. Perhaps buyers who have solely an interaction orientation or solely a self orientation may not be able to survive in an organization if they cannot fulfill task obligations. Given the omnipresence of task orientation, it is not surprising that information exchange resonates with all buyers. Recommendations come close to having universal appeal, with two of three buyer clusters responding to this tactic. Overall, these findings indicate that the quest for unearthing universally effective tactics is not a futile one.

This study also provides insights into tactics that should be avoided. Threats fall into this category. Although theory suggests that threats may be effective with buyers who have a self orientation, we do not find that to be the case. However, two points are noteworthy to put this finding into perspective. First, salespeople may have limited opportunities to use threats because few buyers exhibit a high level of self orientation. Second, the overall use of threats by salespeople in this sample is low, thus raising the possibility that the nonsignificant finding is because of range restriction. A reason for the lack of reliance on threats may have to do with our context. Unlike channels contexts (e.g., Frazier and Summers 1986), in industrial retail settings, dealers and manufacturers often have little power over customers. Any form of threats (e.g., tighter credit terms, delayed delivery) is likely to misfire because customers can turn to other suppliers. Nevertheless, we advocate future studies to retain this tactic because in other selling contexts, customer options may be more limited, and threats may be a viable tactic when dealing with buyers who have a self orientation (McFarland 2003).

Our findings emphasize the importance of the two tactics that we add to Frazier and Summers's (1984) taxonomy and also call into question the comprehensiveness of many sales training programs. We find that ingratiation plays a crucial role in affecting manifest influence of buyers with an interaction orientation and self orientation, whereas inspirational appeals seem to work with buyers with an interaction orientation. With two-thirds of the buyers falling into clusters that exhibit two orientations (i.e., engagement-focus and self-focus-with-balance clusters), these tactics can be used frequently to persuade buyers. Notably, however, several of the best-selling sales training programs in the United States underscore a selling approach that is particularly suited for buyers with a task orientation. The focus of these training programs is on asking a sequence of questions to develop fully the "implications and need payoff" (as in SPIN selling [Rackham 1988]) or in "analyzing current position" and "thinking through alternative positions" (as in Miller, Heiman, and Tuleja's [1985] program strategic selling). Although this is correct given the universal benefit of using information exchange across all clusters of buyers, influence tactics with an affective basis should find a greater voice in the prominent sales training programs.

Overall, when buyers have several orientations, salespeople might be better off using a combination of tactics rather than only one or two tactics. The encouraging news for sales managers is that salespeople's choice of influence tactics appears to be largely governed by buyers' orientations, even when they align along multiple orientations. This suggests that if salespeople can identify buyer cues, they are likely to respond to them.

Limitations and Directions for Further Research

We note the importance of contextual differences. Although the salespeople who participated in this study are geographically dispersed across the United States and work at independently owned dealerships, the salesperson data are nevertheless from a single industry. Thus, there is a need for further research that examines a variety of industries and firms to expand the generalizability of the empirical results. For example, an examination of industries in which buyers are highly dependent on sellers would advance the understanding of the appropriateness and effectiveness of the threats tactic under these circumstances. In the same spirit, it would be important to assess whether the relative importance of the different influence tactics is relatively stable or varies across cultures. A strength of the current study is that though the focus is on professional buyer–seller relationships in a business-to-business context, the salespeople sell both products and services, often in a retail setting.

An important contribution of this study to the influence tactics literature (e.g., DelVecchio et al. 2003; Goff, Bellenger, and Stojack 1994; Scheer and Stern 1992; Spiro and Perreault 1979; Yukl and Tracey 1992) is the validation of the SITs framework. Although we were ultimately able to use existing influence tactics in the formation of this framework, we were careful to follow traditional measurement development standards, including a goal of including the most commonly used tactics in the framework and eliminating those that are not consistent with the personal selling context. For researchers who want to examine the salesperson influence process within the buyer–seller exchange, SITs provide a good starting point. Inspirational appeals have received the least attention in the literature compared with the other five SITs we identified. Furthermore, almost all the literature that has focused on this tactic has not published full item-level data. Thus, four of the five items we included in this study for this measure are new. We recognize that our scale items for this construct will require modifications in future studies. Researchers should develop items that are more specific to the emotions being targeted.[5]

---

[5] We are grateful to a knowledgeable reviewer for pointing out the necessary modifications to the scale items.

APPENDIX
Scale Items

Buyer Responding In Regard to a Specific Salesperson

Source	SITs Scale Items
Based on Venkatesh, Kohli, and Zaltman (1995)	**Information Exchange** The salesperson ... 1. Presented information related to your various purchase options. 2. Asked about your long-term purchasing goals. 3. Ensured you received all product and sales materials relevant to your purchasing decisions.[a, d] 4. Asked if there were any problems or needs he or she could help address.[c, d] 5. Talked about the possible applications of his or her products or services.[d]
Adapted from Venkatesh, Kohli, and Zaltman (1995)	**Recommendations** The salesperson ... 1. Made it clear that by following his or her recommendations, your business would benefit. 2. Outlined the evidence for expecting success from his or her suggestion(s).[c] 3. Provided a clear picture of the positive impact on your business a recommended course of action would have. 4. Made a logical argument supporting his or her specific suggestions.[c, d] 5. Made it explicit, when making a suggestion, that it was intended for the good of your firm.
Adapted from Venkatesh, Kohli, and Zaltman (1995)	**Threats** The salesperson ... 1. Stated that your firm would stop receiving preferential treatment if his or her proposals were ignored. 2. Advised that he or she would stop doing business with your company if his or her requests were not followed. 3. Threatened to become uncooperative if your firm failed to agree to his or her demand(s). 4. Indicated he or she could make "things difficult" for your business if his or her requests were not met.
Adapted from Venkatesh, Kohli, and Zaltman (1995)	**Promises** The salesperson ... 1. Offered to give special attention to your company if you would give him or her new business.[d] 2 Made promises to give something back for complying with his or her request (e.g., discounts, quicker delivery). 3. Offered additional benefits for your business after you had been initially reluctant to agree to his or her terms.[c] 4. Offered to provide incentives to your business for agreeing to his or her purchase request(s). 5. Offered a specific deal for your business to change your position on certain issues.[c]

APPENDIX
Continued

Buyer Responding in Regard to a Specific Salesperson

Source	SITs Scale Items
Adapted from Schriesheim and Hinkin (1990)	**Ingratiation** The salesperson ... 1. Acted in a friendly manner prior to asking for what he or she wanted.[b] 2. Sympathized with you about the added problems that his or her request caused.[a] 3. Made you feel good about yourself before making his or her sales pitch. 4. Complimented and praised your achievements.[d] 5. Discussed shared interests and/or hobbies prior to discussing sales issues.[d]
Based on Yukl and Tracey (1992)	**Inspirational Appeals** The salesperson ... 1. Tried to get you excited about what he or she was selling.[d] 2. Argued you had an exciting opportunity to help your company/business.[a, d] 3. Described the use of his or her products or services with enthusiasm and conviction. 4. Appealed to your values and ideals when asking for your business.[d] 5. Made a sales pitch which tried to appeal to your emotions.[a, d]

Source	Outcome Variable Scale Items
Adapted from Kohli and Zaltman (1988)	**Manifest Influence**[e] 1. How much weight did you give to the salesperson's opinions before buying?[a] 2. To what extent did the salesperson's involvement influence your choices? 3. How much impact did the salesperson have on your purchase decisions? 4. To what extent did you go along with the salespersons suggestions? 5. How much weight did you give the salesperson's statements in making you purchase decisions?[d] 6. To what extent did your decisions reflect the salespersons influence? 7. To what extent did the salesperson influence the criteria used for making purchase decisions?[a]

Salesperson Responding in Regard to Matched Buyer

Source	Buyer Orientation Scale Items
Adapted from Williams and Spiro (1985)	**Buyer Task Orientation**[f] 1. This customer likes to focus on the task at hand. 2. This customer likes to make our sales interactions as efficient as possible. 3. This customer is highly goal oriented.[d] 4. This customer is purposeful in accomplishing his/her purchasing goals.[c]
Adapted from Williams and Spiro (1985)	**Buyer Self Orientation**[f] 1. This customer is more interested in him/herself than what I have to say. 2. This customer usually dominates conversations.[c] 3. This customer tries to impress me with him/herself. 4. This customer is more focused on talking about him/herself than my products.
Adapted from Williams and Spiro (1985)	**Buyer Interaction Orientation**[f] 1. This customer is easy to talk to.[c] 2. This customer likes to socialize during sales interactions. 3. This customer likes to talk to people.[c] 4. This customer is interested in me as a person, not just a salesperson.

---

[a] We dropped this item after the confirmatory factor analysis.

[b] We found this item to have a low factor loading and did not include it in the confirmatory factor analysis.

[c] We dropped this item because of face validity concerns of the reviewers.

[d] We added this item to the original scale.

[e] We measured these items on a five-point Likert-type scale, anchored by "very little" and "a great deal."

[f] We measured these items on a five-point Likert-type scale, anchored by "strongly disagree" and "strongly agree."

Notes: We measured all items on a five-point Likert-type scale, anchored by "never" and "always," except when otherwise noted.

REFERENCES

Achrol, Ravi, Torger Reve, and Louis W. Stern (1983), "The Environment of Marketing Channels: A Framework for Comparative Analysis," *Journal of Marketing*, 47 (Fall), 55–67.

Adaval, Rashmi (2001), "Sometimes It Feels Right: The Differential Weighting of Affect-Consistent and Affect-Inconsistent Product Information," *Journal of Consumer Research*, 28 (June), 1–17.

Ahearne, Michael, Thomas W. Gruen, and Cheryl Burke Jarvis (1999), "If Looks Could Sell: Moderation and Mediation of the Attractiveness Effect on Salesperson Performance," *International Journal of Research in Marketing*, 16 (December), 269–84.

Armstrong, J. Scott and Terry S. Overton (1977), "Estimating Nonresponse Bias in Mail Surveys," *Journal of Marketing Research*, 14 (August), 396–402.

Bass, Bernard M. (1960), *Leadership, Psychology, and Organizational Behavior*. New York: Harper Brothers.

——— (1967), "Social Behavior and the Orientation Inventory: A Review," *Psychological Bulletin*, 68 (October), 260–92.

Bearden, William O., Richard G. Netemeyer, and Jesse E. Teel (1989), "Measurement of Consumer Susceptibility to Interpersonal Influence," *Journal of Consumer Research*, 15 (March), 473–81.

Blake, Robert R. and Jane S. Mouton (1980), *The Grid for Sales Excellence: New Insights into a Proven System of Effective Sales*. New York: McGraw-Hill.

Boyle, Brett F. and F. Robert Dwyer (1995), "Power, Bureaucracy, Influence, and Performance," *Journal of Business Research*, 32 (3), 189–201

———, ———, Robert A. Robicheaux, and James T. Simpson (1992), "Influence Strategies in Marketing Channels: Measures and Use in Different Relationship Structures," *Journal of Marketing Research*, 29 (November), 462–73.

Brown, James R., Robert F. Lusch, and Carolyn Y. Nicholson (1995), "Power and Relationship Commitment: Their Impact on Marketing Channel Member Performance," *Journal of Retailing*, 71 (4), 363–92.

Brown, Steven P. (1990), "Use of Closed Influence Tactics by Salespeople: Incidence and Buyer Attributions," *Journal of Personal Selling & Sales Management*, 10 (Fall), 17–29

———, William L. Cron, and John W. Slocum Jr (1997), "Effects of Goal-Directed Emotions on Salesperson Volitions, Behavior, and Performance: A Longitudinal Study," *Journal of Marketing*, 61 (January), 39–50.

Challagalla, Goutam N. and Tasadduq A. Shervani (1996), "Dimensions and Types of Supervisory Control: Effects on Salesperson Performance and Satisfaction," *Journal of Marketing*, 60 (January), 89–105.

Comstock, Jamie and Gary Higgins (1997), "Appropriate Relational Messages in Direct Selling Interaction: Should Salespeople Adapt to Buyers' Communication Style?" *Journal of Business Communication*, 34 (4), 401–418.

Crosby, Lawrence A , Kenneth R. Evans, and Deborah Cowles (1990), "Relationship Quality in Services Selling: An Interpersonal Influence Perspective," *Journal of Marketing*, 54 (July), 68–81.

DelVecchio, Susan K., James E. Zemanek, Roger P McIntyre, and Reid P. Claxton (2003) "Buyers' Perceptions of Salesperson Tactical Approaches," *Journal of Personal Selling & Sales Management*, 23 (1), 39–49.

Dixon, Andrea L., Rosann L Spiro, and Maqbul Jamil (2001), "Successful and Unsuccessful Sales Calls: Measuring Salesperson Attributions and Behavioral Intentions," *Journal of Marketing*, 65 (July), 64–78.

Doney, Patricia M. and Joseph P. Cannon (1997), "An Examination of the Nature of Trust in Buyer–Seller Relationships," *Journal of Marketing*, 61 (April), 35–51.

Erez, Miriam, Yeshayahu Rim, and Ida Keider (1986), "The Two Sides of the Tactics of Influence: Agent Vs. Target," *Journal of Occupational Psychology*, 59 (March), 25–39.

Evans, F.B. (1963), "Selling as a Dyadic Relationship: A New Approach," *The American Behavioral Scientist*, 6 (May), 76—79.

Fitzsimons, Gavan J. and Donald R. Lehmann (2004), "Reactance to Recommendations: When Unsolicited Advice Yields Contrary Responses," *Marketing Science*, 23 (Winter), 82–94.

Forgas, Joseph P. (1995), "Mood and Judgment: The Affective Infusion Model (AIM)," *Psychological Bulletin*, 117 (1), 39–66.

Fornell, Claes and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 19 (February), 39–50.

Frazier, Gary L. and Raymond C. Rody (1991), "The Use of Influence Strategies in Interfirm Relationships in Industrial Product Channels," *Journal of Marketing*, 55 (January), 52–69.

——— and John Summers (1984), "The Use of Influence Strategies in Interfirm Relationships in Industrial Channels," *Journal of Marketing*, 48 (Summer), 43–55.

——— and ——— (1986), "Perceptions of Interfirm Power and Its Use Within a Franchise Channel of Distribution," *Journal of Marketing Research*, 23 (May), 169–79.

Goff, Brent G., Danny N. Bellenger, and Carrie Stojack (1994), "Cues to Consumer Susceptibility to Salesperson Influence: Implications for Adaptive Retail Selling," *Journal of Personal Selling & Sales Management*, 14 (Spring), 25–39.

Gundlach, Gregory T. and Ernest R. Cadotte (1994), "Exchange Interdependence and Interfirm Interaction: Research in a Simulated Channel Setting," *Journal of Marketing Research*, 31 (November), 516–32.

Hair, Joseph F., Rolph E. Anderson, Ronald L. Tatham, and William C. Black (1998), *Multivariate Data Analysis*, 5th ed. Englewood Cliffs, NJ: Prentice Hall.

Johnston, John (1984), *Econometric Methods*, 3d ed. New York: McGraw-Hill.

Jöreskog, Karl G. and Dag Sörbom (1996), *LISREL 8: User's Reference Guide*. Chicago: Scientific Software International

Keith, Janet, Donald Jackson, and Lawrence Crosby (1990), "Effects of Alternative Types of Influence Strategies Under Different Channel Dependence Structures," *Journal of Marketing*, 54 (July), 30–41.

Kelman, Herbert C. (1958), "Compliance, Identification, and Internalization: Three Processes of Attitude Change," *Journal of Conflict Resolution*, 2 (1), 51–60.

——— (1961), "Process of Opinion Change," *Public Opinion Quarterly*, 25 (Spring), 57–78.

Kim, Keysuk (2000), "On Interfirm Power, Channel Climate, and Solidarity in Industrial Distributor-Supplier Dyads," *Journal of the Academy of Marketing Science*, 28 (3), 388–405.

Kipnis, David and Stuart M. Schmidt (1988), "Upward-Influence Styles: Relationship with Performance Evaluations, Salary, and Stress," *Administrative Science Quarterly*, 33 (December), 528–42.

———, ———, and Ian Wilkinson (1980), "Intraorganizational Influence Tactics. Explorations in Getting One's Way," *Journal of Applied Psychology*, 65 (4), 440–52.

Kohli, Ajay K. (1989), "Effects of Supervisory Behavior: The Role of Individual Differences Among Salespeople," *Journal of Marketing*, 53 (October), 40–50.

——— and Gerald Zaltman (1988), "Measuring Multiple Buying Influences," *Industrial Marketing Management*, 17 (August), 197–204.

Kumar, Kamalesh and Michael Beyerlein (1991), "Construct and Validation of an Instrument for Measuring Ingratiatory Behaviors in Organizational Settings," *Journal of Applied Psychology*, 76 (5), 619–27.

Leigh, Thomas W. and Patrick F. McGraw (1989), "Mapping the Procedural Knowledge of Industrial Sales Personnel: A Script-Theoretic Investigation," *Journal of Marketing*, 53 (January), 16–34.

McFarland, Richard G. (2003), "Crisis of Conscience: The Use of Coercive Sales Tactics and Resultant Felt Stress in the Salesperson," *Journal of Personal Selling & Sales Management*, 23 (Fall), 311–25.

Miles, Morgan P., Danny R. Arnold, and Henry W. Nash (1990), "Adaptive Communication: The Adaptation of the Seller's Interpersonal Style to the Stage of the Dyad's Relationship and the Buyer's Communication Style," *Journal of Personal Selling & Sales Management*, 10 (1), 21–27.

Miller, Robert B., Stephen E. Heiman, and Tad Tuleja (1985), *Strategic Selling: The Unique Sales System Proven Successful by America's Best Companies*. New York: Warner Books.

Milligan, Glenn W. and Martha C. Cooper (1985), "An Examination of Procedures for Determining the Number of Clusters in a Data Set," *Psychometrika*, 50 (2), 159–79.

Netemeyer, Richard G., James S. Boles, Daryl O. McKee, and Robert McMurrian (1997), "An Investigation into the Antecedents of Organizational Citizenship Behaviors in a Personal Selling Context," *Journal of Marketing*, 61 (July), 85–98.

Nunnally, Jum C. (1978), *Psychometric Theory*. New York: McGraw-Hill.

Payan, Janice M. and Richard G. McFarland (2005), "Decomposing Influence Strategies: Argument Structure and Dependence as Determinants of the Effectiveness of Influence Strategies in Gaining Channel Member Compliance," *Journal of Marketing*, 69 (July), 66–79.

Pham, Michel Tuan, Joel B. Cohen, John W. Pracejus, and G. David Hughes (2001), "Affect Monitoring and the Primacy of Feelings in Judgment," *Journal of Consumer Research*, 28 (September), 167–88.

Podsakoff, Philip M., Scott B. MacKenzie, Jeong-Yeon Lee, and Nathan P. Podsakoff (2003), "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies," *Journal of Applied Psychology*, 88 (5), 879–903.

Punj, Girish and David W. Stewart (1983), "Cluster Analysis in Marketing Research: Review and Suggestions for Applications," *Journal of Marketing Research*, 20 (May), 134–48.

Rackham, Neil (1988), *SPIN Selling*. New York: McGraw-Hill.

Ross, William T., Jr., Erin Anderson, and Barton Weitz (1997), "Performance in Principal-Agent Dyads: The Causes and Consequences of Perceived Asymmetry of Commitment to the Relationship," *Management Science*, 43 (May), 680–704.

Scheer, Lisa K. and Louis W. Stern (1992), "The Effect of Influence Type and Performance Outcomes on Attitude Toward the Influencer," *Journal of Marketing Research*, 29 (February), 128–42.

Schriesheim, Chester A. and Timothy R. Hinkin (1990), "Influence Tactics Used by Subordinates: A Theoretical and Empirical Analysis and Refinement of the Kipnis, Schmidt, and Wilkinson Subscales," *Journal of Applied Psychology*, 75 (3), 246–57.

Sheth, Jagdish N. (1976), "Buyer-Seller Interaction: A Conceptual Framework," in *Proceedings of the Association for Consumer Research*. Cincinnati: Association for Consumer Research, 382–86.

Singh, Jagdip and Gary K. Rhoads (1991), "Boundary Role Ambiguity in Marketing-Oriented Positions: A Multidimensional, Multifaceted Operationalization," *Journal of Marketing Research*, 28 (August), 328–38.

Smith, J. Brock and Donald W. Barclay (1997), "The Effects of Organizational Differences and Trust on the Effectiveness of Selling Partner Relationships," *Journal of Marketing*, 61 (July), 3–21.

Spiro, Rosann L. and William D. Perreault Jr. (1979), "Influence Use by Industrial Salesmen: Influence-Strategy Mixes and Situational Determinants," *Journal of Business*, 52 (2), 435–55.

——— and Barton A. Weitz (1990), "Adaptive Selling: Conceptualization, Measurement, and Nomological Validity," *Journal of Marketing Research*, 27 (February), 61–69.

Spitzberg, B.H. and W.R. Cupach (1984), *Interpersonal Communication Competence*. Beverly Hills, CA: Sage Publications.

Stotland, Ezra and Alvin Zander (1958), "The Effect of Public and Private Failure on Self-Evaluation," *Journal of Abnormal and Social Psychology*, 56, 223–29.

Strutton, David, Lou E. Pelton, and John F. Tanner Jr. (1996), "Shall We Gather in the Garden," *Industrial Marketing Management*, 25 (2), 151–62.

Sujan, Harish (1986), "Smarter Versus Harder: An Exploratory Attributional Analysis of Salespeople's Motivations," *Journal of Marketing Research*, 23 (February), 41–49.

Szymanski, David M. (1988), "Determinants of Selling Effectiveness: The Importance of Declarative Knowledge to the Personal Selling Concept," *Journal of Marketing*, 52 (1), 64–77.

Venkatesh, R., Ajay K. Kohli, and Gerald Zaltman (1995), "Influence Strategies in Buying Centers," *Journal of Marketing*, 59 (October), 71–82.

Webster, Frederick E. (1968), "Interpersonal Communication and Salesperson Effectiveness," *Journal of Marketing*, 32 (July), 7–13.

Weitz, Barton A. (1981) "Effectiveness in Sales Interactions: A Contingency Framework," *Journal of Marketing*, 45 (Winter), 85–103.

———, Harish Sujan, and Mita Sujan (1986), "Knowledge, Motivation, and Adaptive Behavior: A Framework for Improving Selling Effectiveness," *Journal of Marketing*, 50 (October), 174–91.

Williams, Kaylene C. and Rosann L. Spiro (1985), "Communication Style in the Salesperson–Customer Dyad," *Journal of Marketing Research*, 22 (November), 434–42.

———, ———, and Leslie M. Fine (1990), "The Customer-Salesperson Dyad: An Interaction/Communication Model and Review," *Journal of Personal Selling & Sales Management*, 10 (Summer), 29–43.

Yukl, Gary and J. Bruce Tracey (1992), "Consequences of Influence Tactics Used with Subordinates, Peers, and the Boss," *Journal of Applied Psychology*, 77 (4), 525–35.

Pierre Chandon & Brian Wansink

How Biased Household Inventory Estimates Distort Shopping and Storage Decisions

The authors develop a model of how consumers estimate the level of product inventory in their households. Two laboratory experiments and two field studies involving 29 product categories show that (1) consumers anchor their estimates on their average inventory and fail to adjust sufficiently; (2) adjustments follow an inelastic psychophysical power function, leading to overestimations of low levels of inventory and underestimations of high levels; and (3) adjustments are more elastic and, thus, more accurate when inventory is salient. Contrary to the assumptions of practitioners and academic modelers, these inventory *estimates*, not actual inventory levels, drive subsequent purchase incidence. Simulation results further show that biased estimates increase overstocking and spoilage among stockout-averse consumers but increase stockouts and unmet demand among overstocking-averse consumers. By predicting the magnitude, not just the direction, of estimation biases, the model and the results offer new insights into accelerating the consumption of healthy foods and improving the targeting of stockpiling-inducing sales promotions.

> "I buy lots of things and then go back to the house and see the fridge is full of all the stuff I've just bought"
> —Prince William, Heir of the British Throne,
> (*BBC News* 2003)

Biased inventory estimates can lead to overstocking and stockouts in a household. Overstocking contributes to households losing an estimated 14% of their meat, grain, fruit, and vegetable purchases (Jones et al. 2003). Household-level stockouts contribute to unmet demand, which represents a wasted opportunity for consumers and for retailers and manufacturers. For example, the success of the "Got milk?" advertising campaign is primarily attributed to its reduction of the frequency of milk stockouts in consumers' refrigerators (Manning 1999). If biased inventory estimations lead to overstocking and stockouts, it is an important issue for the welfare of budget-pressured and time-pressured consumers. It is also an important issue for managers and researchers because these biases can influence storage, consumption, and repurchase decisions and therefore affect promotional elasticity (Meyer and Assunçao 1990). Furthermore, if consumers are biased in their estimations, this would imply that the interpretation and policy recommendations of purchase quantity and timing models may be biased because these models typically assume that consumers know how much of the product is in their pantries (Ailawadi and Neslin 1998).

A recent surge of interest in estimation biases underscores the importance of understanding consumers' inventory estimation biases. Researchers have shown that reference levels, stimulus size, and salience influence consumers' estimations of product volume in a container (Krider, Raghubir, and Krishna 2001; Raghubir and Greenleaf 2006; Raghubir and Krishna 1999), numerosity (Krishna and Raghubir 1997; Pelham, Sumarta, and Myaskovsky 1994), and estimations of purchase and consumption frequency (Lee, Hu, and Toh 2000; Menon, Raghubir, and Schwarz 1995). Still, however, no research has examined the following important questions regarding consumers' inventory estimations that are of interest to consumers, marketers, and researchers: (1) How do people estimate how much product they have in inventory? (2) Are their estimations accurate, or are they systematically biased by the reference inventory, inventory size, and inventory salience? (3) How do biased inventory estimates distort storage and purchase decisions? and (4) What is the relationship between estimation biases and key category characteristics?

To answer these questions, we build on psychophysics research on magnitude estimation and develop a model of how consumers estimate inventory. This model incorporates reference, size, and salience effects. Using the model, we show in a simulation that biased inventory estimates increase overstocking and spoilage among stockout-averse consumers but increase stockouts and unmet demand

---

Pierre Chandon is Assistant Professor of Marketing, INSEAD, (e-mail: pierre.chandon@insead.edu). Brian Wansink is John S. Dyson Chair of Applied Economics and Management, Professor of Marketing, and Professor of Nutritional Science, Department of Applied Economics and Management, Cornell University (e-mail: Wansink@Cornell.edu). The authors thank Gilles Laurent; Collin Payne; Jack Soll; Luc Wathieu; Alex Chernev; Priya Raghubir; Aradhna Krishna; Gita Johar; Sachin Gupta; Gérard Cachon; and those who participated when the authors presented this research at INSEAD; at the University of Illinois, Urbana-Champaign; at the Marketing Science Conference, College Park, Md.; and at the Association for Consumer Research Conference, Toronto.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

among overstocking-averse consumers. We directly test the model's predictions in two laboratory experiments, which include manipulations of internal and external reference levels, of the actual size of the inventory, and of the inventory's salience. Two field studies involving 29 product categories further demonstrate the robustness of the model predictions. They also show that estimated inventory levels predict repurchase decisions better than actual inventory levels do and that inventory estimation biases are related to three key category characteristics: the degree of impulse buying, the ease of stockpiling, and the average promotional elasticity. In the general discussion, we show how our model and the results provide a parsimonious explanation for other, previously unexplained estimation biases. We also offer new insights for accelerating the consumption of healthy foods and improving the targeting of sales promotions

A Model of Consumers' Inventory Estimations

In this section, we develop a model of how consumers estimate the quantity of a product they have in inventory. The key features of the model are that (1) consumers anchor their estimations on internal or external reference levels and insufficiently adjust for the actual inventory level, (2) adjustments are inelastic (they increase more slowly than the true deviation from the reference level, and thus their quality worsens as the deviation increases), and (3) adjustments are more elastic (and, thus, more accurate) when inventory is perceptually salient. We first show how each prediction can be derived from psychophysical research on magnitude estimations and then show how these predictions can be integrated in a simple power model.

Reference Effects

Inventory estimations involve either judgments of numerosity (e.g., "How many eggs do I have?") or judgments of volume (e.g.,"How many ounces of coffee do I have?"). Many studies have shown that consumers anchor numerosity and volume estimations on salient internal or external reference levels and fail to adjust sufficiently for deviations from the reference level. For example, Krishna and Raghubir (1997) show that consumers' estimation of the number of dots in a line is higher when the dots are in multiple clusters (high-reference anchor condition) than when they are all together in one uninterrupted line (low-reference anchor condition). Raghubir and Krishna (1999) show that volume estimations are anchored on the elongation of the container (see also Wansink and Van Ittersum 2003). Krider, Raghubir, and Krishna (2001) show that consumers anchor area estimations on the most salient dimension, in which salience is context dependent. For example, they show that the orientation of a square influences whether the diagonal or a side is used as an anchor when estimating its surface (see also Pelham, Sumarta, and Myaskovsky 1994).

In the context of inventory estimations, we expect that the default anchor is the average category inventory level for each consumer. This is a reasonable assumption because in the absence of other information on actual inventory, the average inventory level is the best estimator of actual inventory if inventory follows a normal or uniform distribution. However, consistent with Krider, Raghubir, and Krishna's (2001) results, we expect that consumers will use external anchors if they are made contextually salient (e.g., by asking consumers to judge explicitly whether an inventory level is above or below some number). In summary, we expect that consumers anchor their inventory estimations on their average inventory, except when external reference levels are made salient (in which case these reference levels serve as anchors).

Size Effects

Recent research on anchoring effects has shown that when people have internally selected a reference as an anchor, they insufficiently adjust for the difference between the reference and the actual value of the magnitude to be estimated (Epley and Gilovich 2001). Epley and Gilovich (2004) show that people estimate the number of days it takes Mars or Neptune to orbit the sun by using the number of days it takes the Earth as an anchor (365 days). As a result, they adjust more for Neptune (mean estimated answer is 3447 days) than for Mars (mean estimated answer is 574 days) because they know that Neptune is further away from the sun than Mars, but they still fall short of the truth (which is 60,225 days for Neptune and 869 days for Mars).

Our model contributes to the literature on anchoring and adjustment by further predicting that the size of the adjustment is inelastic to the actual deviation from the anchor because estimations follow a compressive power function of actual magnitude (i.e., $\text{EST} = a \times (\text{ACT})^b$, where EST is the estimated magnitude, ACT is the actual magnitude, and $b < 1$). In other words, the percentage change in estimations is lower than the percentage change in actual size (Stevens 1986). As a result, adjustments become less effective as the deviation between the reference level and the actual inventory level increases.

There is considerable evidence that magnitude estimations follow a compressive power function of actual magnitudes. For example, Teghtsoonian (1965) finds that the exponent of the power function is approximately .7 when estimating the size of three-dimensional objects. Frayman and Dawson (1981) examine exponents of power functions for different shapes (cubes, spheres, octahedrons, cylinders, tetrahedrons) and find that they are all approximately .6. For perceived numerosity judgments, Krueger (1982) finds a power exponent between .80 and .82. Overall, there is strong support in the literature for our prediction not only that consumers fail to adjust sufficiently for the deviation from the reference level but also that such adjustments are inelastic and therefore increase at a lower rate than the true difference between the reference level and the actual inventory level.

Salience Effects

The power exponent, which measures the elasticity of estimations to actual changes in the magnitude of the stimuli, is influenced by the perceptual salience (e.g., visual prominence) of the different dimensions of the stimulus. Krider, Raghubir, and Krishna (2001) find that the power exponent of area estimations for two-dimensional objects is greater when the salience of secondary dimensions (those that are not used as anchors) is increased. For example, because people anchor area estimations of circles on the length of their horizontal diameters, area estimations and willingness to pay for round pizzas are more sensitive to the actual size of the circle when the vertical diameter is made salient.

Building on these findings, we predict that the elasticity of adjustment improves with the perceptual salience of the product in inventory (and, thus, that the power exponent is greater when inventory is salient than when it is not). Following Krider, Raghubir, and Krishna (2001), we define salience as the ability of the product's inventory to attract attention. For example, the salience of a product's inventory increases if the product is stored in a visible place or if it is purchased or consumed frequently, because these factors increase attention to the actual inventory level. When inventory is salient, consumers are more likely to know whether the reference level they anchor on is wrong and should be adjusted. In the extreme case, consumers may have encoded salient inventory so well that they do not need to rely on the reference level at all. When inventory is not salient, however, consumers may not even know whether actual inventory is above or below the reference level and therefore will rely mostly on the reference anchor. Thus, we expect that inventory estimations are adjusted slightly when inventory is not salient and are adjusted more strongly when inventory is salient. In other words, we expect that inventory estimations are more sensitive to actual inventory and, thus, more accurate when inventory is salient than when it is not.

Modeling Biases in Inventory Estimations and Their Effects on Storage and Shopping Decisions

A Model of Household Inventory Estimation

We develop a simple mathematical model of consumers' inventory estimations that allows us to test the effects of inventory reference, size, and salience simultaneously. This model applies to situations in which consumers cannot simply look and count their inventory and, therefore, must estimate it from memory (e.g., when they are in a store). In essence, we estimate a series of power models in different conditions (e.g., when salience is low versus high) and use the model parameters to test for the hypothesized effects. The general form of these power models is as follows:

$$(1) \qquad \text{EST} = a \times (\text{ACT})^{b},$$

where EST is estimated inventory, ACT is actual inventory, and a and b are parameters estimated with regression. We use power models because they are consistent with psychophysical research on magnitude estimation and because they have four desirable properties.

The first desirable property is that the model intercept captures systematic differences between estimated and actual inventory, regardless of inventory level. Changes in the intercepts shift the power curve up or down but leave the shape of the curve unchanged. Therefore, we can test our prediction that internal or external anchors shift estimations toward the reference level by examining the value of the intercept a across anchor conditions.

Second, the power exponent b measures the elasticity of estimations and influences the shape of the power curve. If $b = 1$, estimations increase at the same rate as actual inventory. If $b < 1$, the power function is compressive, and estimations are inelastic (i.e., the percentage change in estimations is lower than the percentage change in actual inventory, or in other words, estimations increase at a slower rate than actual inventory). For example, if $b = .5$ and actual inventory increases by 50% (i.e., by a factor of 1.5), estimations increase by a factor of $(1.5)^{.5} = 1.225$ (i.e., 22.5%). If $b > 1$, the power function is expansive, and estimations are elastic (they increase at a faster rate than actual inventory). Therefore, we can test our prediction that inventory estimations are inelastic by verifying that $b < 1$.

Third, we can use the power model to compute the crossover inventory level at which estimations are accurate ($\text{ACT}^* = a^{1/(1-b)}$). Inventory levels below ACT* tend to be overestimated, and inventory levels above ACT* tend to be underestimated. We can then compare ACT* with the hypothesized reference level (e.g., a person's average perceived inventory).

Fourth, the power function treats positive and negative deviations from the crossover level similarly (i.e., the power exponent is the same, regardless of inventory level; there is no asymmetry between "gains" and "losses"). When the relevant logarithmic measures of accuracy are used (e.g., the log ratio of estimated divided by actual inventory), the magnitude of the bias is a linear function of actual inventory (measured in logs), and its slope is the same above and below the crossover level. Note, however, that the symmetry in the power model leads to asymmetry when more conventional measures of accuracy are used, such as absolute error (AE; |estimated – actual|) or percentage absolute error (PAE; |estimated – actual|/actual). A positive deviation from the crossover inventory level always leads to a smaller PAE (but to a larger AE) than a similar-size negative deviation from the crossover inventory level. For example, if $\text{EST} = 3 \times \text{ACT}^{.5}$, $\text{ACT}^* = 9$ units. When inventory is above the crossover level by 6 units (ACT = 15 units), the AE and PAE are 3.38 units and 23%, respectively. When inventory is below the crossover level by 6 units (ACT = 3 units), the AE and PAE are 2.2 units and 73%, respectively.

To test the hypothesized reference and salience effects, we estimate a different power model for each reference and salience condition. We illustrate how salience effects can be incorporated, but the approach is the same when we examine the effects of different inventory reference levels:

$$(2) \qquad \text{EST} = a \times (\text{ACT})^{b}, \text{ when salience is low, and}$$

$$\text{EST} = a' \times (\text{ACT})^{b'}, \text{ when salience is high.}$$

To facilitate the statistical estimation of these power models, we linearized them as follows:[1]

---

[1] The linearization prevents the model from being estimated when either estimated magnitude or actual magnitude is zero. To overcome this limitation, it is possible to estimate the following nonlinear model with the Levenberg–Marquardt least square algo-

$$\text{(3)}\quad \ln(\text{EST}) = \ln(a) + b \times \ln(\text{ACT}),\ \text{when salience is low, and}$$

$$\ln(\text{EST}) = \ln(a') + b' \times \ln(\text{ACT}),\ \text{when salience is high.}$$

To test whether a is statistically different from a′ and whether b is statistically different from b′, we estimate the following moderated regression:

$$\text{(4)}\qquad \ln(\text{EST}) = \alpha + \beta \times \text{SAL} + \delta \times \ln(\text{ACT}) + \gamma \times \ln(\text{ACT}) \times \text{SAL} + \varepsilon,$$

where SAL (salience) is a binary variable coded as $-\frac{1}{2}$ when salience is low and as $\frac{1}{2}$ when salience is high; ε is the error term; and α, β, δ, and γ are estimated with ordinary least squares (OLS). These parameters are interpreted as follows: α measures the average intercept across the two salience conditions ($e^{\alpha} = [a \times a']^{\frac{1}{2}}$ is the geometric mean of the two intercepts); β measures the main effect of salience on the intercepts ($e^{\beta} = a'/a$); δ measures the average power exponent across the two salience conditions ($\delta = [b + b']/2$); and γ measures the effects of salience on the power exponent, that is, the interaction between salience and actual inventory ($\gamma = b' - b$). If salience improves the accuracy of estimations by reducing the degree of compression, we would expect γ to be positive and statistically different from 0.

Simulating the Effects of Biased Inventory Estimations on Storage and Shopping Decisions

Because no one expects that consumers have perfect inventory knowledge, the important issue is to determine the extent to which these biases distort consumers' storage and purchase decisions. To address this issue, we simulate the weekly category purchase decisions of consumers for a perishable product with stochastic household-level demand and examine the conditions under which biased inventory estimations lead to overstocking and, thus, to spoilage or stockouts and, ultimately, to unmet demand.

Consider a shopper who must decide whether to buy a dozen eggs in each weekly shopping trip but does not know what the exact demand will be over the next days because the three household members have an independent 50% chance of wanting an egg for breakfast. On any given day, demand can be 0 ($p = 1/8$), 1 ($p = 3/8$), 2 ($p = 3/8$), or 3 ($p = 1/8$) eggs. For the sake of simplicity, we consider two different types of households. The first segment consists of overstocking-averse households that discard their eggs quickly (e.g., after 9 days because they consume their eggs soft boiled and are afraid of bacteria). The optimal purchasing rule for overstocking-averse households is to buy a dozen eggs if their inventory on the day of their weekly shopping trip falls below 6 eggs. The second segment consists of stockout-averse households that keep their eggs for a relatively long time (e.g., 14 days, because they are consuming them hard-boiled). The optimal purchasing rule for stockout-averse households is to buy a dozen eggs if their inventory is below 12 eggs.

Our model of inventory estimation allows us to simulate how biased inventory estimates influence the magnitude of spoilage and unmet demand for households that are averse either to stockouts (with a discard threshold of 14 days and a purchasing threshold of 12 eggs) or to overstocking (discard threshold = 9 days, purchasing threshold = 6 eggs). Building on the field study results for eggs (see Field Study 2), we model inventory estimations for eggs as a compressive power function with parameters $a = b = 1$ (i.e., EST = ACT) when estimations are unbiased or with parameters $a = 3$ and $b = .5$ (i.e., $\text{EST} = 3 \times [\text{ACT}]^{.5}$) when estimations are biased.

For stockout-averse consumers, because of the underestimation of high inventory levels, when their estimate is 12 eggs, the actual inventory is 16 eggs (because $3 \times [16]^{.5} = 12$). As a result, stockout-averse consumers continue to repurchase a dozen egg even if they have as many as 16 eggs in inventory, whereas the optimal purchase threshold for them is 12 eggs. Figure 1 illustrates the effects of estimation biases for these households in terms of spoilage and unmet demand over the first eight weeks of the year. Over the course of one year, Table 1 shows that estimation biases cause stockout-averse consumers to purchase more (+7%, or 36 eggs), to increase the amount of spoilage strongly (+225%, or 27 eggs), and to reduce unmet demand only slightly (–33%, or 2 eggs).

For overstocking-averse consumers, because of the overestimation of low inventory levels, when their estimate is 6 eggs, the actual inventory is 4 eggs (because $3 \times [4]^{.5} = 6$). As a result, overstocking-averse consumers forgo repurchasing a dozen eggs even when they have as few as 4 eggs in inventory, whereas the optimal purchase threshold for them is 6 eggs. Figure 2 illustrates the effects of estimation biases for these households in terms of spoilage and unmet demand over the first eight weeks of the year. Over the course of one year, Table 1 shows that estimation biases cause overstocking-averse consumers to buy less (–9%, or 48 eggs), to increase unmet demand strongly (+118%, or 52 eggs), and to reduce spoilage slightly (–23%, or 12 eggs).

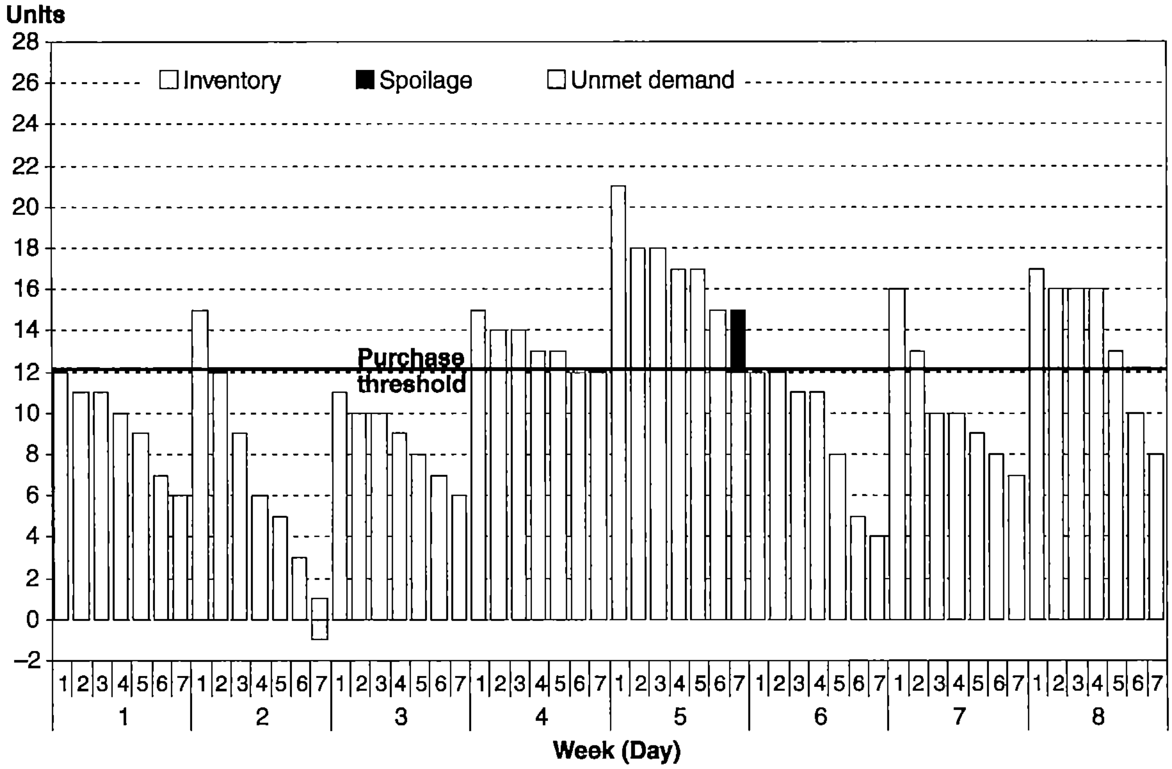
Overall, the simulation shows that biased inventory estimations can significantly increase both spoilage and unmet demand, depending on whether consumers are averse to stockouts or to overstocking. The underestimation of high inventory levels causes stockout-averse households to repurchase when their inventory is too high, thus leading to a significant increase in spoilage. Conversely, the overestimation of low inventory levels prevents overstocking-averse consumers from repurchasing when their inventory is too low, thus causing a significant increase in unmet demand. In addition, Table 1 shows that biased inventory estimates increase the variance of inventory and increase shopping and storage inefficiencies (the sum of units spoiled because of overstocking and those not consumed because of stockouts) for all households. In the "General Discussion" section, we show how these effects provide marketers with new opportunities for increasing consumption and for improving the targeting of sales promotions.

---

rithm: $\text{EST} = \alpha \times (\beta^{\text{SAL}}) \times (\text{ACT})^{(\delta \times (\gamma)^{\text{SAL}})}$, where $\alpha = (a \times a')^{\frac{1}{2}}$, $\beta = a'/a$, $\delta = (b \times b')^{\frac{1}{2}}$, and $\gamma = b'/b$. The interpretations of the α, β, δ, and γ parameters is the same as in the linearized model. In all analyses, both methods provided similar estimates, and therefore we use the simpler linearized model, except when incorporating zeroes is desirable because of a low number of observations.

FIGURE 1
How Biased Inventory Estimates Influence Daily Inventory, Spoilage, and Unmet Demand for Stockout-Averse Households (Purchase Threshold = 12 Units)

A: With Unbiased Inventory Estimates



B: With Biased Inventory Estimates

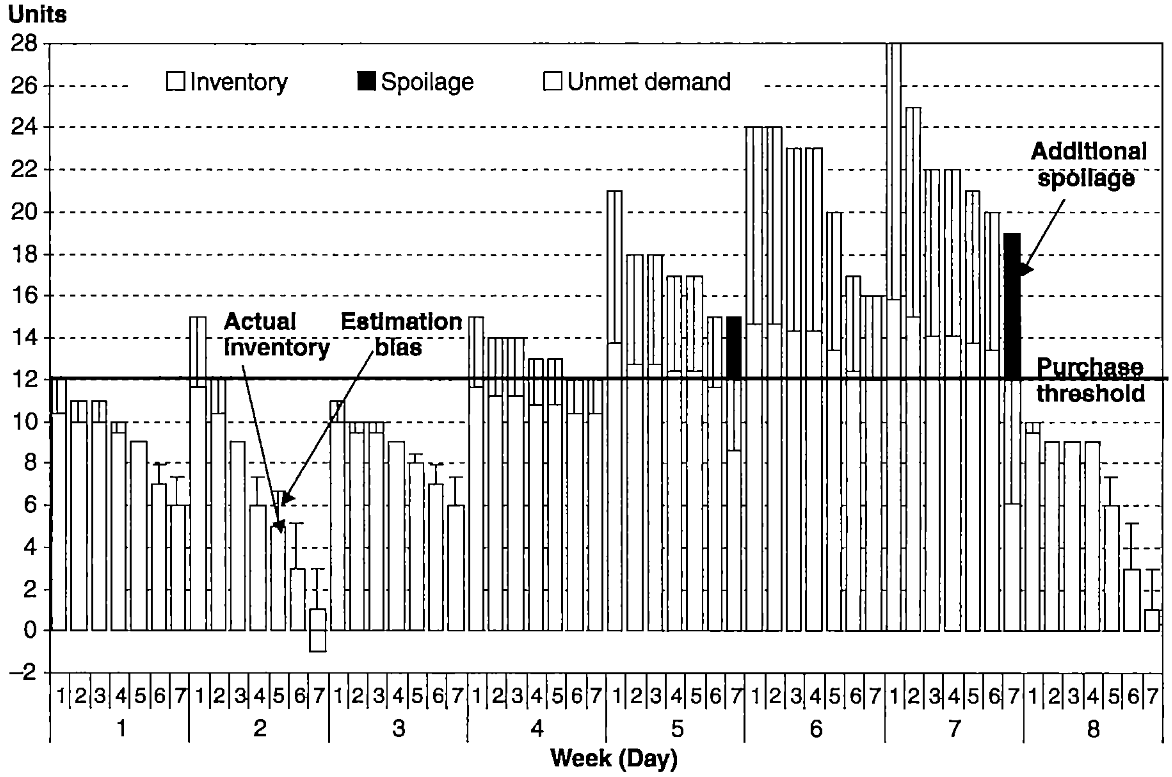


TABLE 1
Simulation Results: How Biased Inventory Estimates Influence Purchase Quantity, Spoilage, Unmet Demand, and Daily Inventory

	Stockout-Averse Households (Purchase Threshold = 12 Eggs)			Overstocking-Averse Households (Purchase Threshold = 6 Eggs)		
Inventory Estimations	**Unbiased (Actual PT = 12 Eggs)**	**Biased (Actual PT = 16 Eggs)**	**Difference**	**Unbiased (Actual PT = 6 Eggs)**	**Biased (Actual PT = 4 Eggs)**	**Difference**
Demand	539	539	0	539	539	0
Purchase quantity	552	588	+7%	552	504	–9%
Spoilage	12	39	+225%	53	41	–23%
Unmet demand	6	4	–33%	44	96	+118%
Spoilage plus unmet demand	18	43	+139%	97	137	+41%
Daily inventory (M)	11.1	14.0	+27%	7.9	6.6	–16%
Daily inventory (SD)	4.8	5.4	+13%	4.1	4.3	+4%

Notes: Spoilage is the annual number of eggs discarded. The discard threshold is 14 days for stockout-averse consumers and 9 days for overstocking-averse consumers. Unmet demand is the number of eggs that were not consumed because of stockouts. The simulation assumes that consumers purchase a dozen egg if their inventory estimate on the day of their weekly shopping trip is below the purchase threshold (PT). Because of the underestimation of high inventory levels, the actual purchase threshold for biased stockout-averse households is too high (16 versus 12 eggs). Because of the overestimation of low inventory levels, the actual purchase threshold for biased overstocking-averse households is too low (4 versus 6 eggs).

Experiment 1: How Inventory Size and External Anchors Influence Inventory Estimations

Procedure

The objective of Experiment 1 is to test how external anchors and inventory size influence inventory estimations (we test the effects of internal anchors and of inventory salience in Experiment 2). To achieve this objective, Experiment 1 used a mixed design with one within-subject factor with four levels (one, three, seven, or nine units in inventory), one between-subject factor with three levels (no external anchor, low external anchor, or high external anchor), and two replications (two different products) for each level of inventory.

Participants first examined a color picture of a pantry containing 8 target products and 13 other products in different quantities. The pantry contained one unit of 2 target products, three units of 2 other target products, seven units of 2 other target products, and nine units of 2 other target products. Participants were later asked to recall the inventory level of the 8 target products (Coca-Cola cans, Lifesavers candy, Smucker's jam, Campbell's soup, Charmin toilet tissue, Crest toothpaste, Carr's crackers, and Heinz tomato sauce). We rotated the products across the four inventory-level conditions following a Latin-square design. We selected the inventory levels of the target products on the basis of a prestudy of 21 consumers, who indicated that this range (one to nine) would straddle the average inventory of these products for almost all respondents (the typical average inventory for these products is three or four units).

The participants were 216 undergraduate students who were awarded extra credit participation points for a course they were attending. We told them that the aim of the study was to measure their liking of different types of teas. Consistent with this, we first asked them to evaluate the three brands of tea that were present in the pantry. To direct their attention toward the other products in the pantry (including the eight target products), we asked the participants to estimate the overall quality of the brands displayed and to indicate whether some of these products should be stored in a refrigerator rather than a pantry. We then asked each participant to return the first booklet containing the pantry picture. After a brief distracter task, we gave participants a second booklet containing a typical anchoring manipulation. We asked participants in the high-external-anchor condition (low-external-anchor condition) to indicate whether the number of units of each of the eight target products was above or below nine (one). We then asked them to estimate the number of units of the target products. Participants in the control condition estimated the number of units of the target products. Finally, we asked all the participants to write down how they had estimated the inventory levels for these products.

Results

We consider first the estimations made in the control (no external anchor) condition. Figure 3 shows that the mean estimated inventory was well below reality when there were seven or nine units in inventory, close to reality when there were three products in inventory, and slightly above reality when there was only one product in inventory. To test size effects in the control condition formally, we estimated the following linear regression with OLS:

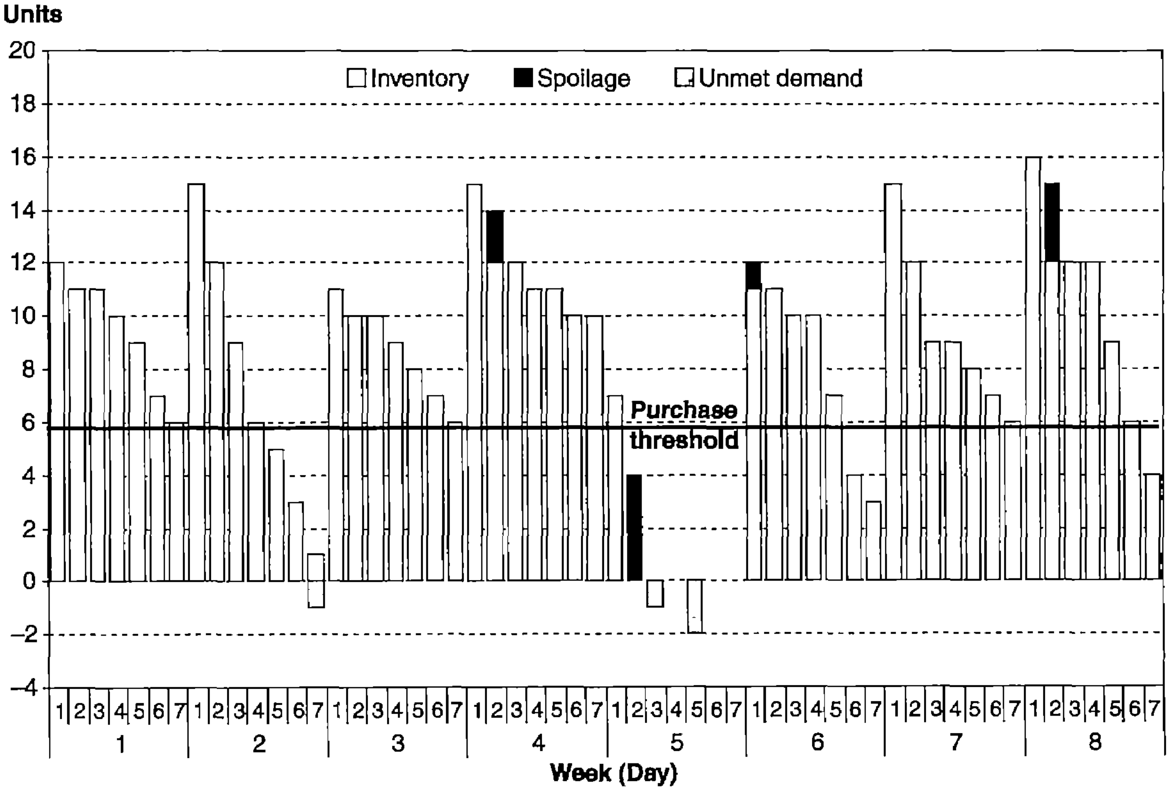
$$\ln(EST_{ij}) = \alpha + \beta \times \ln(ACT_{ij}) + \Sigma\sigma_j \times CAT_j + \varepsilon, \tag{5}$$

where EST_{ij} is the estimated inventory of product j by participant i, ACT_{ij} is the actual inventory, CAT_j are seven binary variables accounting for product differences, and ε is the error term.

As we expected, the power exponent was statistically below 1 ($\alpha = b = .43$, t-test of difference from 1 = –9.1, $p <$

FIGURE 2
How Biased Inventory Estimates Influence Daily Inventory, Spoilage, and Unmet Demand for Overstocking-Averse Households (Purchase Threshold = 6 Units)

A: With Unbiased Inventory Estimates



B: With Biased Inventory Estimates

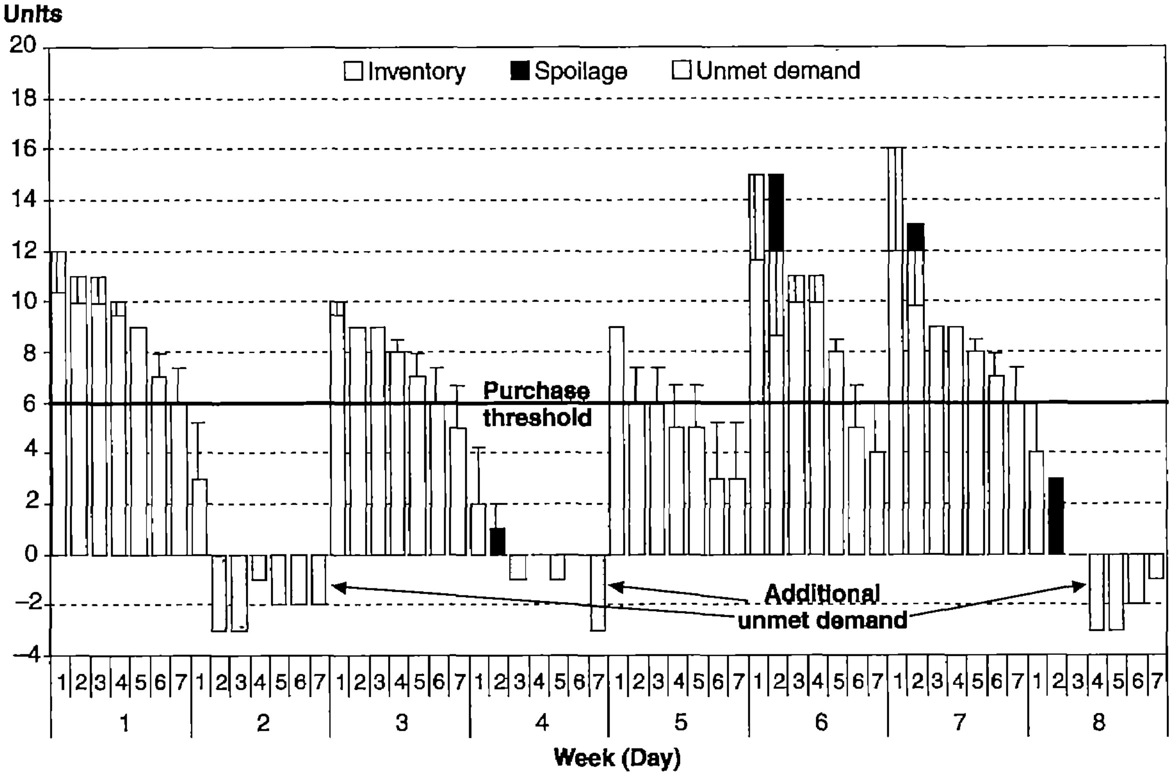
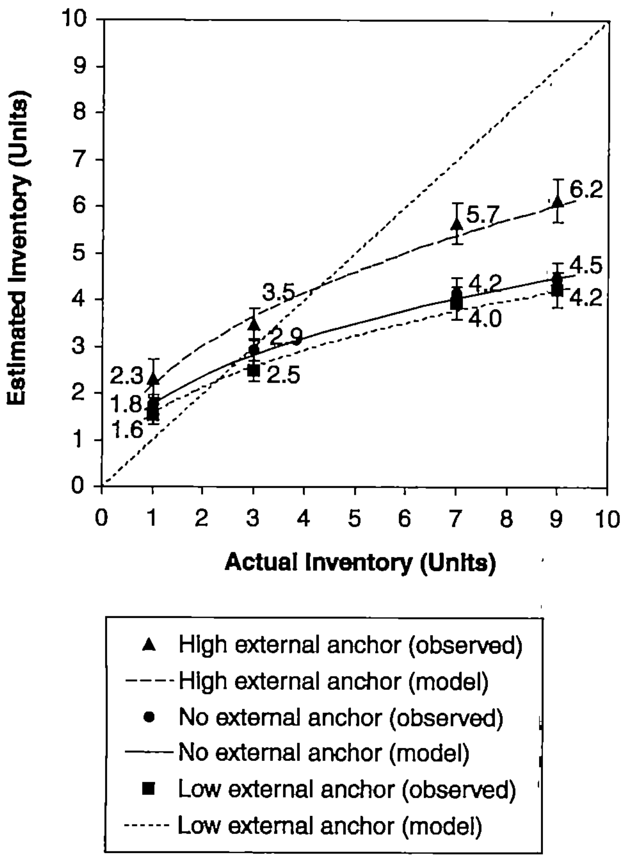


FIGURE 3
Experiment 1: Effects of External Anchors on Inventory Estimations (Geometric Means, Confidence Intervals, and Model Predictions)



.001), and the intercept was statistically above 0 ($\alpha = \ln(a) = .58$, $t = 11.8$, $p < .001$), which indicates that the intercept of the power model ($a = 1.79$) is statistically larger than 1. Five of the seven product-specific intercepts were statistically significant. In addition, we compared the fit of the power model with that of an alternative linear model ($EST_{ij} = \alpha + \beta \times ACT_{ij} + \Sigma\delta_j \times CAT_j + \varepsilon$). As we expected, the R-square was stronger for the power model ($R^2 = .36$) than for the linear model ($R^2 = .33$), and the mean average percentage error (MAPE) was lower for the power model (MAPE = .58) than for the linear model (MAPE = .74; paired t-test = 21.3, $p < .001$). In summary, as we predicted, consumer inventory estimations in the control condition follow a compressive power function of actual inventory.

We now turn to the analysis of the anchoring manipulation. As a manipulation check, we asked two coders, both of whom were unaware of the objective of the experiment, to classify consumers' retrospective protocols into three categories. The "internal anchor" category included protocols that referred to how much of the product a person usually has in inventory. The "visual memory" category included protocols that recalled the picture of the pantry. The "other" category included the remaining protocols, such as "I just guessed." Consistent with the anchoring literature, which shows that people are unaware of the effects of external anchors (Mussweiler, Strack, and Pfeiffer 2000), no protocol mentioned the anchoring manipulation. An analysis of the 166 usable protocols shows that the frequency of protocols mentioning internal anchors was higher in the control condition (M = 22.1%) than in the two external anchor conditions (M = 7.1%; $\chi^2 = 7.8$, $p < .01$). This suggests that, as we predicted, consumers were more likely to use internal anchors when external anchors were not contextually salient.

Figure 3 shows that the anchoring manipulation shifted inventory estimations toward the anchor but did not influence the relationship between estimated and actual inventory (the power curves are parallel). To test anchoring effects formally, we estimated the following regression:

$$
\begin{aligned}
(6)\quad \ln(EST_{ij}) = {} & \alpha + \beta \times \ln(ACT_{ij}) \\
& + \delta \times EXTANCH1_i + \gamma \times EXTANCH9_i \\
& + \lambda \times EXTANCH1_i \times \ln(ACT_{ij}) \\
& + \theta \times EXTANCH9_i \times \ln(ACT_{ij}) + \Sigma\sigma_j \times CAT_j + \varepsilon_{ij},
\end{aligned}
$$

where EST_{ij} is the estimated inventory for product j by participant i, ACT_{ij} is the (geometric) mean-centered actual inventory for product j and participant i,[2] $EXTANCH1_i$ is a binary variable taking the value of ⅔ if participant i was in the low-external-anchor condition (anchor = 1) and –⅓ otherwise, $EXTANCH9_i$ is a binary variable taking the value of ⅔ if participant i was in the high-external-anchor condition (anchor = 9) and –⅓ otherwise, and CAT_j are seven binary variables accounting for product differences (j = [1, ..., 7]). The simple effects of both external anchors were statistically significant and in the expected direction (for the anchor on one unit, $\delta = -.07$, $t = -2.0$, $p < .05$; for the anchor on nine units, $\gamma = .27$, $t = 7.8$, $p < .01$), indicating that inventory estimations were assimilated toward the anchors. Consistent with the model, the interaction parameters were not statistically significant ($\lambda = .03$, $t = .6$, $p = .52$, and $\theta = .2$, $t = .4$, $p = .68$), indicating that the degree of compression of inventory estimations was the same across the three conditions. Finally, the intercept was statistically different from 0 ($\alpha = 1.21$, $t = 78$, $p < .01$). This indicates that participants significantly overestimated their inventory when it was only one unit (in the control condition). Four of these product-specific intercepts were statistically significant.

---

[2] In practice, we divide actual inventory by its mean across inventory conditions (five units). In a moderated regression, mean centering the variables involved in an interaction allows us to estimate the main effects of the variables involved in an interaction when the other variables involved in the interaction are at their mean. For example, it allows us to estimate the effects of anchoring on inventory estimations when inventory level is at its average level (five units). We use geometric mean centering rather than arithmetic mean centering to be consistent with the power functional form and to prevent negative values whose logs cannot be computed.

Discussion

Experiment 1 shows that inventory estimations are assimilated toward external anchors and are adjusted for the actual inventory level through a nonlinear compressive power function. Experiment 1 also shows that the rate of adjustment for the actual size of the inventory remains constant, regardless of which reference levels serve as anchors. In other words, external anchors do not influence the exponent of the power function and thus do not influence the quality of adjustments from the reference level. Finally, the protocols provide indirect evidence that consumers rely on internal anchors when external anchors are not salient. However, because we did not measure the value of these internal anchors, we cannot test the effects of internal anchors or determine whether an average home inventory can serve as one. Experiment 2 further tests the model by examining the effects of internal anchors and by directly manipulating product salience.

Experiment 2: How Internal Anchors, Inventory Salience, and Inventory Size Influence Inventory Estimations

Procedure

Experiment 2 used the same procedure and stimuli as Experiment 1 but with three important differences. First, we did not manipulate external anchors but asked each participant to indicate the average inventory of the eight target products in their own house (the hypothesized internal anchor). Second, we manipulated the perceptual salience of the target products in three combined ways. Salient products were located on the top or middle shelf of the pantry (as opposed to the bottom shelf), were separate from other products (rather than being crowded together with them), and were given multiple facings when available in more than one unit (rather than being stacked together in an overlapping fashion). We assigned each of the eight target products to one of the eight conditions created by the two (high or low salience) × four (one, three, seven, or seven units) design. As in Experiment 1, we rotated the products across the eight inventory size and salience conditions following a Latin-square design. Third, as a means to check the effectiveness of the salience manipulation, we asked participants to rate how visible each product was in the picture.

The participants were 150 undergraduate students who were awarded extra credit participation points for a course they were attending. Manipulation checks show that the salience manipulation was successful. An analysis of variance indicated that products in the high-salience condition were rated as "more visible" (M = 6.75 on a nine-point scale, anchored by 1 = "completely disagree" and 9 = "completely agree") than those in the low-salience condition (M = 5.89; F(1, 1090) = 24.6, $p < .001$). However, neither the inventory size nor the salience manipulation influenced the home inventory reported for each product (F(3, 1090) = .58, $p = .63$, and F(1, 1090) = .59, $p = .44$, respectively). This shows that the average home inventory level was not estimated from the inventory level or the salience of the product in the study.

Results

Figure 4, Panel A, shows the mean estimated inventory as a function of the inventory and salience manipulations. As we expected, increasing salience made estimations less compressive and more accurate. To examine the effects of internal anchors, we assigned participants to a high- or low-internal-anchor group on the basis of their self-reported average home inventory level for each product. Across the eight products, the average home inventory in the high-internal-anchor group was M = 8.5 units versus M = .9 units in the low-internal-anchor group. Figure 4, Panel B, shows mean estimated inventory as a function of actual inventory for both internal anchor groups. As we expected, estimations in the high-internal-anchor condition were higher than those in the low-internal-anchor condition, regardless of inventory level, but the two estimation curves remained parallel, suggesting that internal anchors (unlike salience) did not influence the elasticity of adjustments.

To test our predictions that salience reduces the degree of compression of estimations and that estimations shift toward internal reference levels, we estimated the following regression:

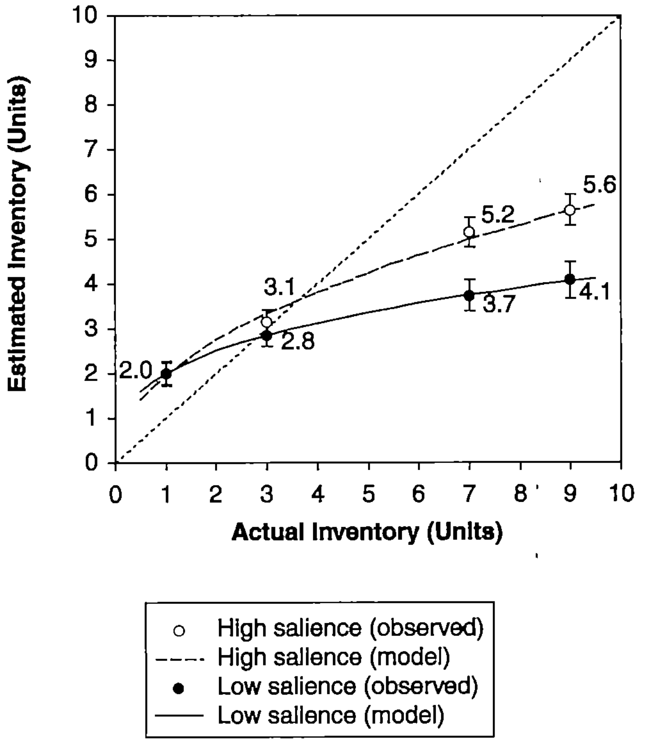
$$\begin{aligned}\text{(7)}\quad \ln(EST_{ij}) &= \alpha + \beta \times \ln(ACT_{ij}) + \delta \times INTANCH_{ij}\\ &+ \gamma \times SAL_{ij} + \lambda \times INTANCH_{ij} \times \ln(ACT_{ij})\\ &+ \theta \times SAL_{ij} \times \ln(ACT_{ij}) + \Sigma\sigma_j \times CAT_j + \varepsilon_{ij},\end{aligned}$$

where EST_{ij} is the estimated inventory for product j by participant i, ACT_{ij} is the geometric mean-centered actual inventory for product j and participant i, $INTANCH_{ij}$ is a binary variable taking the value of ½ if the home inventory of participant i for product j is in the top 50% of the distribution for this product and –½ otherwise, SAL_{ij} is a binary variable taking the value of ½ if product j is in the high-salience condition for participant i and –½ otherwise, and CAT_j are seven binary variables accounting for product differences (j = [1, ..., 7]).

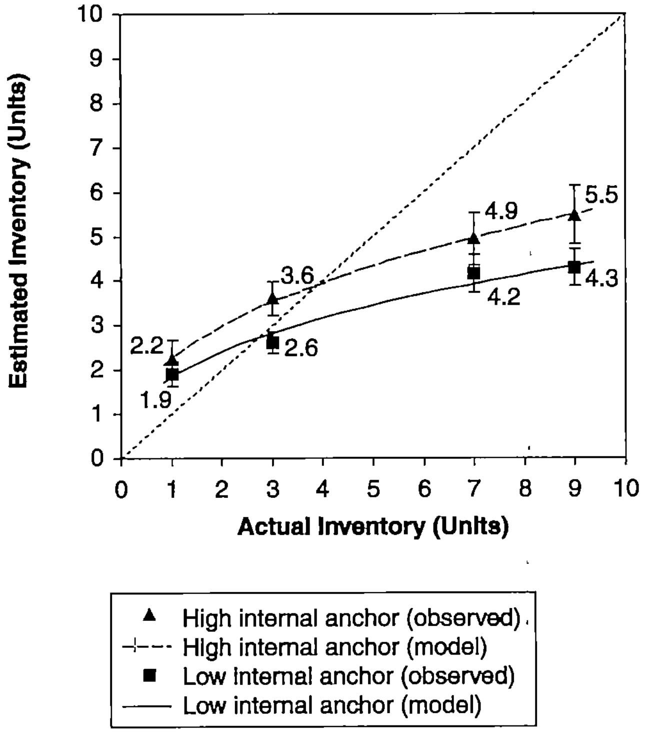
As in Experiment 1, the power exponent was statistically below 1 ($\beta = .41$, t-test of difference from 1 = –28.9, $p < .001$), indicating that the rate of adjustment increases more slowly than the actual inventory size. The coefficient capturing the simple effect of the internal anchor was positive and statistically significant ($\delta = .09$, t = 2.3, $p < .05$), but its interaction with the actual inventory level was not ($\lambda = .02$, t = .5, $p = .61$). This shows that inventory estimations were assimilated toward the average home inventory for that product but that this internal anchor did not change the rate at which estimations were adjusted for the actual inventory level. The main effect of the salience manipulation was positive and statistically significant ($\gamma = .22$, t = 6.5, $p < .01$), and its interaction with the actual inventory level was also positive and statistically significant ($\theta = .15$, t = 3.7, $p < .01$). Because of the significant interaction between the actual inventory level and salience, the effects of salience are not statistically significant when the inventory level is low (one or three units). This shows that adjustments are more sensitive when inventory is salient than

FIGURE 4
Experiment 2: Effects of Perceptual Salience and Internal Anchors on Inventory Estimations (Geometric Means, Confidence Intervals, and Model Predictions)

A: Perceptual Salience



B: Internal Anchors



when it is not. Finally, the intercept was statistically different from 0 ($\alpha = 1.23$, $t = 71.3$, $p < .01$), and four of the seven product-specific intercepts were statistically significant, replicating the results of Experiment 1.

Discussion

Experiment 2 shows that internal anchors, like the external ones, shift estimations toward the reference level but do not change the rate at which estimations are adjusted for the actual inventory level. Experiment 2 also shows that the salience of the product in the pantry increases the rate of adjustment. Estimations of less salient products are almost entirely driven by the reference inventory (the power curve is almost flat). In contrast, estimations of salient products are significantly influenced by the actual inventory level (the slope of the curve is close to one).

Taken together, Experiments 1 and 2 provide strong evidence in support of the model of inventory estimations. Yet although judgment and estimation biases are often found in a laboratory setting, they can be less apparent, or even negated, in the field, in which there is less variation in inventory size, reference levels, and salience and consumers may have greater experience with the estimation task. In Field Study 1, we investigate the robustness of these effects by measuring actual and estimated inventory and product salience in six categories. We also study whether estimated inventory is a better predictor of category purchase incidence than actual inventory. In Field Study 2, we test the robustness of the model in 23 new categories, and we examine whether the degree of compression of a given category is associated with the degree of impulse buying, the ease of stockpiling, and the promotional elasticity of that category.

Field Study 1: Inventory Salience and Estimation Accuracy

Procedure

Over two periods of five days each, we intercepted 121 adult consumers in four different central Illinois supermarket parking lots as they were exiting the supermarket, and we asked them to estimate their current inventory of six product categories in exchange for $9. After completing their estimates, participants were given a preaddressed, stamped envelope and a brief questionnaire that asked them to check their actual inventory levels for these categories when they returned home the same day. The same questionnaire also asked them to rate the visibility of each category in their homes by indicating their agreement with the following sentence: "These [category name] are stored in a very visible place." They rated their agreement on a nine-point scale, anchored by 1 = "strongly disagree" and 9 = "strongly agree." Finally, participants indicated whether they had purchased any products from these six categories during their latest shopping trip.

We chose the six categories (apples, canned tuna, hot dogs, potatoes, tea bags, and tomatoes) on the basis of a pretest involving 16 consumers, which showed that they estimated their inventory using discrete package units (as opposed to continuous measures, such as ounces) and that there was enough variance in the way different households stored them to expect salience effects (e.g., some consumers keep tomatoes in a salient container on the kitchen counter, whereas others store them in the refrigerator where they are less visible). To verify the quality of the measures of actual inventory levels, we visited 16 households the day after they sent us their pantry checks and measured the number of units for these six categories ourselves. When we excluded partial units and subsequently purchased units, pantry accuracy was well over 90% for this subsample of households. Of the 121 consumers intercepted, 90 (74.4%) returned their questionnaire within one week, and we included them in the analysis.

Results

Purchase incidence. Part of our motivation for studying biases in inventory estimations is the assumption that estimations, not actual inventory, drive important decisions, such as whether to repurchase from a given category during a supermarket shopping trip. Field Study 1 enables us to test this assumption by comparing the association between category purchase incidence on the one hand and estimated or actual inventory on the other hand.

We measured purchase incidence using a variable ($REFILL_{ij}$) that took the value of 1 if at least one purchase was made from category j by participant i during the supermarket shopping trip that occurred just before the estimation and 0 if otherwise. As we expected, the correlation between purchase incidence and estimated inventory was negative and statistically significant ($r[REFILL_{ij}, EST_{ij}] = -.10$, $p < .05$), whereas the correlation between purchase incidence and actual inventory was not statistically different from zero ($r[REFILL_{ij}, ACT_{ij}] = -.03$, $p = .52$). To test whether the predictive power of estimated and actual inventory is statistically different, we used a repeated measures analysis of variance with EST_{ij} and ACT_{ij} as the within-subjects measures and $REFILL_{ij}$ as the between-subjects factors. As we expected, the interaction between $REFILL_{ij}$ and the within-subjects factor (estimated versus actual inventory) was statistically significant ($F(1, 441) = 3.8$, $p < .05$).[3] This result implies that purchase incidence is more strongly associated with estmated inventory than with actual inventory.

Inventory estimation biases. We found that the mean estimated inventory was within 10% of the mean actual inventory level for the six categories. However, this aggregate accuracy is not a result of consumers being accurate. Rather, it is a result of underestimations compensating for overestimations. Across the six categories, only 49% of inventory estimations were accurate, whereas 28% were underestimations and 23% were overestimations. For example, Figure 5 shows that the mean inventory estimation of tea bags and tomatoes is slightly overestimated at low levels (the two lowest quartiles) but is strongly underestimated at high levels (the two highest quartiles). Figure 5 also shows that these estimates were less compressive and, thus, more accurate when tomatoes and tea bags were highly salient than when they were not.

To test the statistical significance of size and salience directly, we estimated the following regression:

$$\ln(EST_{ij}) = \alpha + \beta \times \ln(ACT_{ij}) + \delta \times SAL_{ij} + \gamma \times SAL_{ij} \times \ln(ACT_{ij}) + \Sigma\sigma_j \times CAT_j + \varepsilon_{ij}, \tag{8}$$

where EST_{ij} is the estimated inventory for category j by participant i, ACT_{ij} is the actual inventory for category j and participant i (it is not mean centered so as to be able to test the intercept when the actual inventory is equal to 1), SAL_{ij} is a mean-centered binary variable measuring the visibility of category j in the pantry of participant i (it is categorized using a median split, but the results are unchanged if we use the original continuous measure), and CAT_j are five binary variables accounting for product differences ($j = [1, \ldots, 5]$). As we expected, the intercept was statistically larger than zero ($\alpha = .39$, $t = 6.2$, $p < .01$), indicating that the intercept of the power model is larger than 1 ($a = e^{\alpha} = 1.48$). Consistent with the hypothesized size effects, the exponent was statistically lower than 1 ($b = \beta = .77$, t-test of difference from $1 = -7.5$, $p < .01$), indicating that inventory estimations are also compressive in the field. The main effect of salience was not statistically significant ($\delta = -.09$, $t = -.9$, $p = .36$), which means that as in Experiment 2, salience did not influence estimations when there was only one unit in inventory. Consistent with the hypothesized salience effects, the interaction between salience and actual inventory was positive and statistically significant ($\gamma = .08$, $t = 2.0$, $p < .05$). As in Experiment 2, inventory estimations were more compressive when the category was not particularly visible in the pantry, and they were more accurate when the category was visible in the pantry.

Discussion

Field Study 1 makes two contributions. First, it shows that estimated inventory is more strongly related to category purchase incidence than is actual inventory. Importantly, this result cannot be explained by mere-measurement effects, because inventory estimations were measured as consumers were exiting the store, *after* their purchases had been made. In addition, our decision to measure purchase incidence with the second questionnaire (completed at home) and not during the parking lot interview (when inventory estimations were collected) reduces the likelihood that these results are driven by self-presentation biases (i.e., that consumers reported inventory estimations consistent with their purchase decisions).

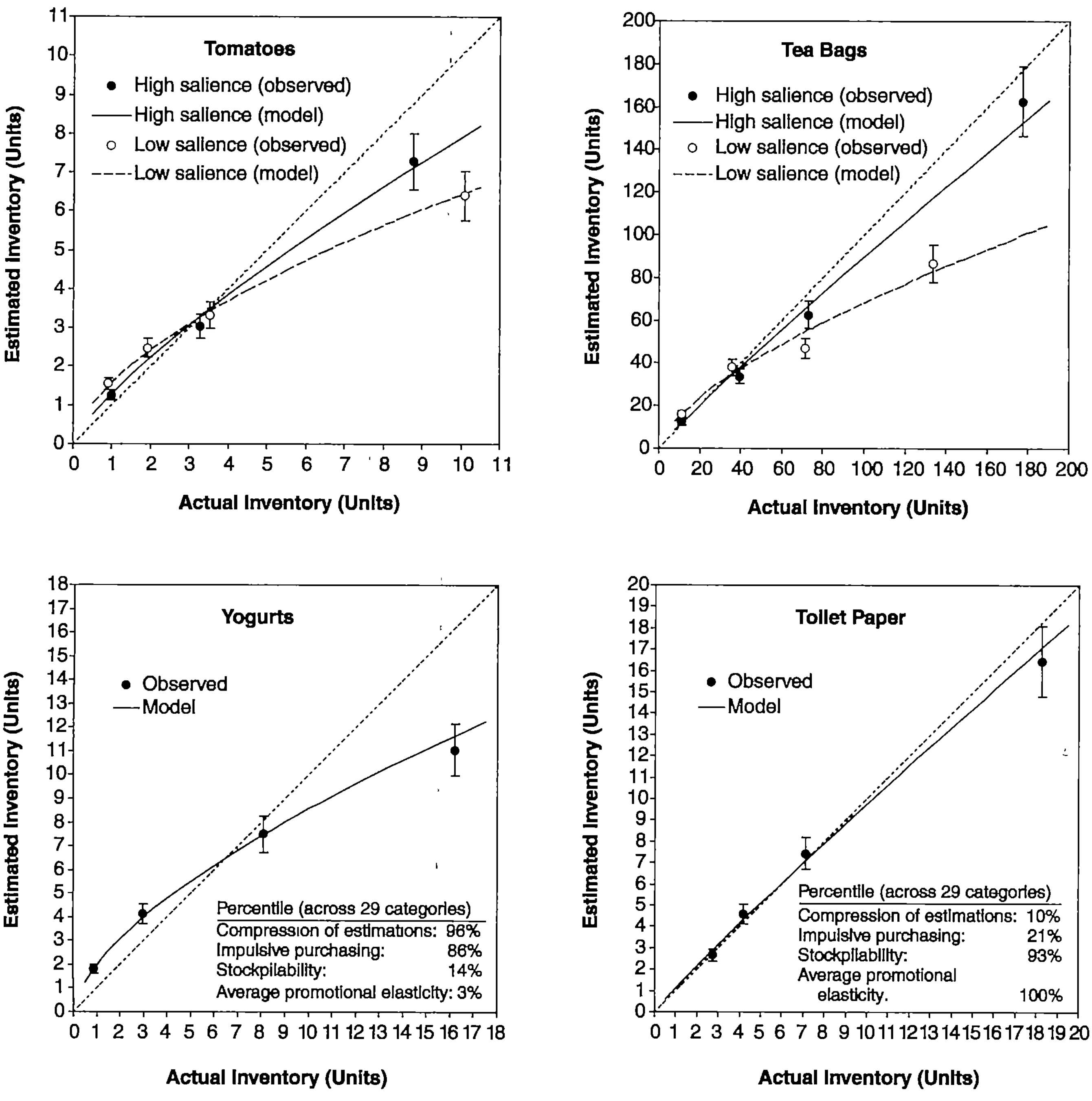
Second, Field Study 1 supports the experimental results by showing that inventory estimations follow a compressive function of actual inventory levels, even when they are made by adult consumers for products that they have just bought. In addition, it shows that consumers adjust for

---

[3] The main effect of the within-subjects factor was not statistically significant ($F(1, 441) = 1.6$, $p = .20$), indicating that mean estimated inventory is similar to mean actual inventory. The main effect of $REFILL_{ij}$ was also not statistically significant ($F(1, 441) = 2.0$, $p = .16$), which is not surprising, given that it measures the average effect of a nonsignificant association (between $REFILL_{ij}$ and ACT_{ij}) and of a significant but weak association (between $REFILL_{ij}$ and EST_{ij}).

FIGURE 5

Field Studies 1 and 2: Estimated and Actual Inventory for Selected Product Categories (Geometric Means, Confidence Intervals, and Model Predictions)



Notes: The degree of compression of inventory estimations of the product category is based on the estimated power exponent in the psychophysical equation EST = a × $(ACT)^b$, the degree of impulse purchasing and the stockpilability of the category are based on Narasimhan, Neslin, and Sen (1996), and the average promotional elasticity was measured by Information Resources Inc (*P-O-P Times* 1991).

actual inventory levels more when the products are salient (stored in a visible place) than when they are less salient.

A limitation of Field Study 1 is that it did not measure the average product inventory for each household. Therefore, its findings should be interpreted as showing that respondents with lower actual inventory tend to overestimate their inventory, whereas respondents with higher inventory levels tend to underestimate them. Naturally, the average inventory across respondents may not be representative of the average inventory of either very light or very heavy buyers. However, note that this heterogeneity in average inventory level reduces the likelihood of detecting inventory biases in the population (because extreme inventory levels are actually average levels for these extreme consumers).

Field Study 1 raises three further questions: (1) Are these results generalizable to nonfood products? (2) Do average inventory levels serve as estimation anchors in the field? and (3) Is the degree of compression of inventory estimations related to managerially important category characteristics? We address the first question by conducting a large-scale field study in which we measure estimated and actual inventory for 23 new food and nonfood categories. We address the second question by incorporating data on average inventory level (rather than on current actual inventory level), which we collected in an additional survey. Finally, we address the third question by studying the association between the power exponent, which measures the degree of compression in each category, and three category characteristics, which we obtained from secondary data, namely, the degree of impulse buying in the product category, the ease of stockpiling the product, and the average promotional brand elasticity of the brands in the product category.

Field Study 2: How Estimation Biases Vary Across Products

Procedure

Field Study 2 used the same procedure as Field Study 1 to measure biases in inventory estimations. In Field Study 2, however, we measured inventory levels in ounces rather than in units for seven products (soft drinks, coffee, shampoo, mayonnaise, laundry detergent, dishwashing detergent, and ketchup) for which inventory levels are typically measured in ounces because of the large variations in package sizes. We measured inventory levels for the other 16 products (soap, canned soup, spaghetti, vacuum cleaner bags, yogurt, toothpaste, frozen meat, eggs, frozen vegetables, butter sticks, canned fruit, pasta sauce, cookies, toilet tissue, salad dressing, and breakfast cereals) in units. Another difference is that we did not measure inventory salience.

To avoid respondent fatigue, we surveyed participants on five to eight products. To verify the accuracy of the actual inventory measures, we asked a subgroup of consumers to telephone one of the researchers immediately after they had checked their actual inventory levels. We called the consumers who had not telephoned by 7:30 that evening and reminded them to check their inventory. There were no systematic differences among the results of participants who had telephoned, the results of participants who needed to be reminded, and the rest of the participants who were not contacted. To check accuracy further, we told an additional subgroup of consumers to keep their questionnaire because we would pick them up the next day. During the pickup round, we requested permission to inspect their actual inventory. With the households who agreed to the inspection (approximately 75%), there were no full-unit discrepancies with the self-reported inventory. Of the 461 consumers who participated in Field Study 2, 317 (68.7%) returned their questionnaire in a timely manner. Together with data we obtained in Field Study 1, we had a total of 2185 estimations on 29 products (an average of 75 observations per category).

To avoid common method biases, we obtained data on impulse buying, ease of stockpiling, and promotional response from two independent sources. For impulse buying and ease of stockpiling, we used the survey of 108 product categories that Narasimhan, Neslin, and Sen (1996) conducted and published in *Journal of Marketing*.[4] This survey measured impulse buying by asking 100 consumers to rate their agreement with the following two statements: "I often buy the product on a whim when I pass by it in the store," and "I typically like to buy this product when the urge strikes me." We measured ease of stockpiling by asking the same consumers to rate their agreement with the following two statements: "It is easy to store extra quantities of this product in my home," and "I like to stock up on this product when I can." For promotional elasticity, we used data from the "Infoscan Topical Marketing Report," generated by Information Resources Inc. and published in the *P-O-P Times* (1991). This report provides an estimate of the average percentage brand sales increase in response to a 15% price cut with an in-store display for 164 product categories based on the results of Information Resources Inc.'s PromotionScan model (Abraham and Lodish 1993) estimated on the checkout data of 2400 grocery stores. The category definitions of Narasimhan, Neslin, and Sen's (1996) survey and of the Infoscan Report matched ours in 25 of the 29 product categories. We used data from categories closest to the remaining four (e.g., frozen side dishes for frozen vegetables).

Results

Are these results generalizable to nonfood products? As in Field Study 1, we found that the mean estimated inventory is similar to the mean actual inventory. With four exceptions (crackers, ketchup, tomatoes, and potatoes), the mean estimations are all within 20% of the mean actual inventory for the 29 products (the 6 products surveyed in Field Study 1 and the 23 surveyed in Field Study 2). As in Field Study 1, this aggregate accuracy does not indicate that consumers are accurate at the individual level; instead, it indicates that overestimations compensate for underestimations. Across the 29 products, only 41% of inventory estimations are accurate (within one unit or one ounce of the actual inventory), whereas 31% are underestimations, and 25% are overestimations.

To test the model predictions, we estimated the power model shown in Equation 1 for each of the 29 categories. Because of the low number of observations for some categories, we used the nonlinear Levenberg–Marquardt least square algorithm, which allows us to incorporate observations with a zero-estimated or actual inventory level (we obtained similar results when estimating a linearized model on nonzero observations with OLS).

---

[4] Scott Neslin, Albert Wesley Frey Professor of Marketing in the Tuck School of Business at Dartmouth College, generously provided these data.

Table 2 shows that the power model fits the inventory estimation data well (the mean R^2 is 54%). All power exponents are below 1, and the variance from 1 is statistically significant at the 5% level for all 29 products. All power intercepts are above 1, and the difference is statistically significant at the 5% level for 22 of the 29 products and is statistically significant at the 10% level for 3 other products. In addition, the fit of the power model was superior to the fit of a linear model (EST = a + b × ACT) for almost all the products. Specifically, the R-square of the power model was higher than the R-square of the linear model for 23 of the 29 products, and the average fit improvement was 3% and was statistically significant. Similarly, the MAPE was lower for the power model than for the linear model for 28 of the 29 products and was statistically lower for the power model than for the linear model across all observations ($MAPE_{power} = .44$ versus $MAPE_{linear} = .61$; paired $t = 10.5$, $p < .001$). All these results support our prediction that inventory estimations for these 29 products follow a compressive power function of actual inventory level.

Using the estimated model parameters, we computed the crossover inventory level ($e^{\alpha/(1 - \beta)}$) for each product. As we expected, the crossover inventory level was within the range of observed actual inventory levels for all 29 products. Excluding one outlier (tea bags, which has high inventory levels when measured in units), we found that low inventory levels tended to be overestimated but that as actual inventory reached four to six units or 24–42 ounces, estimations tended to become accurate. However, when actual inventory levels were above these average levels, they tended to be strongly underestimated.

Do average inventory levels serve as anchors in field estimations? We asked 37 adult consumers, similar to those involved in Field Studies 1 and 2, to estimate their *average* inventory level (as opposed to their current inventory level) for these 29 products. We then estimated the following constrained model:

$$EST_{ij}/AVG_j = \alpha \times (ACT_{ij}/AVG_j)^{\beta} + \varepsilon_{ij}, \tag{9}$$

TABLE 2
Field Studies 1 and 2: Category-Level Power Regression Results (Estimates, Standard Errors, Fit, and Predicted Crossover Inventory Level)

Category	Intercept		Exponent		r(EST, ACT)	N	ACT*
Apples	1.23**	(.13)	.83**	(.04)	.93	89	3.4
Butter sticks	2.03*	(.68)	.63**	(.14)	.63	51	6.8
Canned fruit	1.52**	(.25)	.71**	(.08)	.86	52	4.2
Canned soup	1.81**	(.22)	.69**	(.06)	.78	155	6.8
Canned tuna	1.61**	(.22)	.66**	(.07)	.78	98	4.1
Cereals	1.35**	(.16)	.77**	(.06)	.88	53	3.7
Coffee[a]	2.61**	(.92)	.73**	(.09)	.71	56	34.9
Cookies	1.46**	(.13)	.42**	(.06)	.75	53	2.0
Dishwashing detergent[a]	3.40*	(1.60)	.61**	(.11)	.60	55	23.1
Eggs	3.08**	(1.09)	.54**	(.14)	.61	33	11.5
Frozen meat	2.34**	(.71)	.53**	(.13)	.61	39	6.1
Frozen vegetables	1.74**	(.28)	.61**	(.08)	.80	52	4.1
Hotdogs	2.45**	(.72)	.57**	(.09)	.68	87	8.0
Ketchup[a]	3.08**	(.91)	.57**	(.08)	.69	55	13.7
Laundry detergent[a]	3.51*	(1.71)	.71**	(.09)	.67	54	75.9
Mayonnaise[a]	3.75**	(1.27)	.60**	(.09)	.73	56	27.2
Pasta sauce	1.14	(.12)	.84**	(.08)	.85	53	2.3
Potatoes	1.52	(.47)	.78**	(.10)	.74	88	6.7
Salad dressing	1.36**	(.23)	.66**	(.11)	.70	53	2.5
Shampoo[a]	4.87**	(2.13)	.55**	(.11)	.55	57	33.7
Soap	1.79**	(.21)	.68**	(.06)	.74	147	6.2
Soft drinks[a]	14.36	(11.64)	.41**	(.16)	.50	55	91.5
Spaghetti	1.43**	(.11)	.53**	(.05)	.69	159	2.1
Tea bags	2.07**	(.61)	.82**	(.05)	.83	85	56.9
Toilet paper	1.57**	(.24)	.81**	(.05)	.91	52	10.7
Tomatoes	1.35	(.29)	.69**	(.10)	.71	85	2.6
Toothpaste	1.36**	(.12)	.75**	(.06)	.89	42	3.4
Vacuum cleaning bags	1.78**	(.26)	.59**	(.06)	.66	140	4.1
Yogurt	3.29**	(.60)	.42**	(.07)	.66	131	7.8

*Statistically different from 1 at the 10% level (one-tailed).

**Statistically different from 1 at the 5% level (one-tailed).

[a] Inventory measured in ounces.

Notes: EST = estimated inventory, ACT = actual inventory, N = number of estimations, and ACT* = crossover inventory (mean estimated inventory is equal to mean actual inventory).

where EST_{ij} is the estimated inventory for product j by participant i at the time of the study, ACT_{ij} is the actual inventory, and AVG_j is the average inventory for product j measured in the additional survey. This model is a simple rewriting of the basic power model in Equation 1, in which estimated and actual inventory are expressed as a proportion of the average inventory of the product rather than as the original units. This transformation influences the intercept but leaves the power exponent and the model fit unchanged from the values shown in Table 2.

If people use the average inventory level as an anchor, inventory estimations are unbiased when actual inventory is at its average level, and therefore the mean estimated inventory is equal to the mean actual inventory level when actual inventory is at its average level. In Equation 9, this occurs only if the intercept (α) is equal to 1 (because if $ACT_{ij} = AVG_j$, $EST_{ij} = AVG_j = ACT_{ij}$). As expected, we found that the intercept of this new model was not statistically different from 1 in 25 of the 29 categories. We obtained similar results when using the mean actual inventory level, which we computed across participants in Field Studies 1 and 2 as the category average inventory rather than the mean obtained from the new survey. Overall, these results are consistent with the hypothesis that average inventory levels serve as internal anchors in the absence of salient external anchors.

Is the degree of compression of inventory estimations related to managerially important category characteristics? As we show in Table 2 and Figure 5, some products, such as yogurt, exhibit strong compression and therefore are somewhat inelastic to actual changes in inventory. For example, given that the power exponent of yogurt is .42, if its inventory increases by 50%, estimations increase by only 19% (see Figure 5). Figure 5 also shows that other products, such as toilet paper, exhibit little compression and therefore are relatively accurate at all levels of inventory. Because the power exponent of toilet paper is .81, if its inventory increases by 50%, estimations increase by almost the same percentage (39%). We expect that these category differences in the rate of adjustment to actual inventory are linked to three key category characteristics: (1) the likelihood of impulse buying, (2) the ability to stockpile, and (3) the average brand promotional elasticity in that category.

The actual inventory of products that are likely to be bought on impulse is apt to fluctuate more and in a less predictable way than the actual inventory of products whose purchases are planned. As a result, we expect that estimations will be more compressive and, therefore, less accurate for products with a high degree of impulse purchasing. In contrast, the actual inventory of products that are easy to stockpile should be easier to monitor than the inventory of products that are difficult to stockpile. We also expect that estimations will be less compressive for products that are easy to stockpile. Finally, consumers are more likely to switch to another brand or to stockpile in response to a promotional display when they have an accurate understanding of their inventory than when they do not know how much of the product is left in their inventory and therefore hesitate to buy the promoted brand (Wansink, Kent, and Hoch 1998). Consumers who have no idea about how much of the product they have in inventory are more likely to pass a promotion and follow their habitual purchasing pattern for fear of overstocking. Consequently, we expect estimations to be more compressive for categories with a high average promotional elasticity.

We tested these hypotheses by estimating the following model with the nonlinear Levenberg–Marquardt least square algorithm:

$$EST_{ij} = \alpha \times \beta_1^{IMP_j} \times \beta_2^{STO_j} \times \beta_3^{PROM_j} \times (ACT_{ij})^{\delta} \times \gamma_1^{IMP_j} \times \gamma_2^{STO_j} \times \gamma_3^{PROM_j}, \tag{10}$$

where EST_{ij} is the estimated inventory for category j by participant i, ACT_{ij} is the actual inventory for category j and participant i, IMP_j is a mean-centered binary variable measuring the degree of impulse buying in the category, STO_j is a mean-centered binary variable measuring the ease of stockpiling the product category, and $PROM_j$ is a mean-centered binary variable measuring the average promotional elasticity in the category (we dichotomized the three binary variables with a median split).

As we expected, the parameter capturing the interaction between ACT and IMP was statistically below 1 ($\gamma_1 = .93$, asymptotic 95% confidence interval [CI] = [.85, .99]), indicating a lower power exponent (i.e., less accuracy) for categories bought on impulse. Conversely, the parameter capturing the interaction between ACT and STO was statistically above 1 ($\gamma_2 = 1.15$, 95% CI = [1.04, 1.26]), indicating a higher power exponent (i.e., more accuracy) for categories that are easy to stockpile. Finally, the parameter capturing the interaction between ACT and PROM was above 1 but was statistically significant only at the 10% level ($\gamma_3 = 1.07$, 90% CI = [1.00, 1.15]), indicating a slightly higher power exponent (i.e., more accuracy) for categories with high average promotional elasticity. In summary, these analyses supported our hypotheses that inventory estimates are particularly biased for product categories that are bought on impulse, are difficult to stockpile, and have a low average promotional elasticity.

General Discussion

Consumers' biases in inventory estimations can distort storage and shopping decisions and can lead to overstocking and spoilage or to stockouts and unment demand. To understand the origin of these biases, we developed a model of how consumers estimate the amount of product they have in inventory. This model predicts that (1) unless an external reference level is salient, consumers anchor their estimations on their average inventory level and insufficiently adjust for the actual inventory level; (2) the adjustment from the average inventory follows an inelastic power function; and (3) the adjustment is more elastic and, thus, more accurate when inventory is perceptually salient. Drawing on this model, we simulated the effects of biased inventory estimates on the amount of spoilage and unmet demand for consumers who are averse to stockouts or to overstocking. We then tested the predictions of the model for 8 products in two laboratory experiments and for 29 products in two field studies.

The key results of these studies are as follows: First, contrary to the assumptions of practitioners and modelers, consumers' estimates of their inventory are better predictors of category purchase than their actual inventory. Second, although individual inventory estimations are seldom accurate, the mean estimated inventory of a product category is a valid estimation of its mean actual inventory because underestimations tend to compensate overestimations. Third, in the absence of salient external reference levels, consumers anchor their inventory estimations on their average inventory level and adjust insufficiently for the actual size of the inventory. Fourth, adjustments from the reference level follow an inelastic psychophysical function (a power function with an exponent below 1), and as a result, the quality of the adjustments deteriorates as the inventory deviates from the reference level. In short, below-average inventory levels are slightly underestimated, average inventory levels are accurately estimated, and above-average inventory levels are strongly underestimated. Fifth, inventory estimations are more elastic (more sensitive to actual changes in inventory) and, thus, more accurate when inventory is salient than when it is not. Sixth, the underestimation of high inventory levels causes stockout-averse households to repurchase when their inventory is too high, thus leading to a significant increase in spoilage. Conversely, the overestimation of low inventory levels prevents overstocking-averse consumers from repurchasing when their inventory is too low, thus causing a significant increase in unmet demand. Finally, the least elastic and, thus, least accurate estimations are those of product categories that are bought on impulse, are difficult to stockpile, and have a low promotional elasticity.

Implications for Further Research

By identifying three possible sources of biases, our model contributes to the literature on estimations, which has focused primarily on documenting biases rather than on explaining why these biases occur and what factors moderate them. For example, despite the considerable evidence that people underestimate the duration of past and future events when their duration is greater than five minutes and overestimate them when they last less than five minutes, there is no consensus about why this occurs (Lee, Hu, and Toh 2000; Roy, Christenfeld, and McKenzie 2005). Still, this pattern of results, even in this seemingly unrelated context, can be readily explained by two features of our model: (1) People anchor estimations on average duration, and (2) adjustments for deviations from the average are inelastic.

For researchers interested in magnitude estimations, our basic finding that deviations from the reference level follow a compressive power function has important implications. First, it enables them to predict the magnitude, not just the direction, of these biases. Second, it enables them to predict the average estimation bias for any duration, not just for those already measured. Third, it can lead to new testable predictions, such as the magnitude of estimation biases will become larger as the duration increases, even when the magnitude of the bias is measured as a percentage deviation from reality (for a proof of this, see Chandon and Wansink, in press). Fourth, this model can enable researchers to make predictions about the boundary conditions of duration estimation biases. For example, it would predict that duration estimations are more accurate when the event is salient and when average duration is used as an anchor than when the event is not salient or when external reference levels are used as anchors.

Our finding that inventory estimations follow a compressive power function of actual inventory is consistent with psychophysics research and with a great deal of accumulated evidence on magnitude estimation studies but is the opposite of what signal detection theory would predict. A key feature of signal detection theory is its assumption that people account for the relative costs of over- and underestimations (Green and Swets 1988). For inventory estimations, overestimations are more costly when inventory is low and when stockouts are likely to occur. In contrast, underestimations are more costly when inventory is high and when overstocking is likely. As a result, signal detection theory would predict that consumers underestimate low inventory levels (to avoid costly overestimations) and overestimate high inventory levels (to avoid costly underestimations). One explanation for our opposite findings is that in our studies, the cost–benefit payoff of the estimations was constant across inventory levels. Further research could try to reconcile the psychophysics and signal detection predictions by manipulating the costs of over- and underestimations (e.g., by measuring or manipulating the opportunity costs of having too much or too little). Another explanation is that wishful thinking leads consumers to bias their estimations optimistically in the direction of what they hope they will be rather than in the direction that will help them avoid overstocking and stockouts. Collecting process data, especially over time, may help in teasing out these explanations and in further advancing the understanding of the origins of inventory estimation biases.

Finally, our findings challenge the interpretation of the results of purchase quantity and timing models, which assume that consumers have accurate, or at least unbiased, knowledge of how much of a given product they have in inventory (Ailawadi and Neslin 1998; Bell, Chiang, and Padmanabhan 1999; Sun 2005). These models are estimated on scanner panel data, which contain no information on inventory. Therefore, they estimate inventory from a person's purchase quantity and timing data. Recall our finding that estimated inventory predicts repurchase decisions better than actual inventory. This indicates that the inventory estimated by these models may be what consumers estimate the inventory to be and not what it really is. Thus, these models measure the effects on purchases of consumers' inventory estimations rather than the effects of their actual inventory. Our finding that inventory estimations are inelastic implies that these models overestimate the effects of actual inventory on purchase decisions. Larger changes in actual inventory (and, thus, deeper price cuts) may be necessary to achieve the effect sizes reported in these studies. Biases in inventory decisions could easily be incorporated into purchase timing and quantity models. One solution would be to incorporate a power model of inventory estimation within these models. Another would be to convert the actual inventory levels predicted by these models directly into inventory estimates using the relevant parameters estimated in the field studies and shown in Table 2.

Implications for Managers and Consumers

Our model and the results offer new insights into accelerating the consumption of healthful foods. An important finding of prior research on inventory was the "stock pressure effect"; that is, inventory stockpiling accelerated the consumption of foods (Chandon and Wansink 2002). Unfortunately, it was convenient snack foods that received the greatest boost in consumption frequency when stockpiled. Given the rising trend in obesity, responsible managers of healthful foods could use our new insights into how consumers estimate inventory to influence their consumption of these foods. One way, as was done with the "Got milk?" campaign, would be to raise inventory salience. This would reduce the likelihood of wasting healthful foods, which are often more perishable than less healthful ones. It would also accelerate the consumption rate of health foods. Studies of food-intake diaries have shown that the intake frequency of fruits and yogurts increases as their perceived expiration date approaches (Wansink 2006). Such efforts can also lead people to think of a recent instance in which they consumed the food, which in turns increases consumption intentions (Wansink and Deshpandé 1994). Because inventory estimations are anchored on a person's own typical or normal level of inventory, another way of increasing a person's purchase and consumption of a healthful food would be to increase his or her perception of what should be a typical or normal level of inventory of that food. For example, knowing that the average family of four has 11.2 cans of soup in inventory could lead consumers to alter their inventory (and consumption) norm for the relatively healthful option of canned soup (Wansink and Chandon 2006).

Our model and the results also offer new insights into the targeting of sales promotions to consumer and category segments. Consider the two segments of consumers shown in Table 1 with identical demand: overstocking-averse consumers (who only repurchase when their inventory is low) and stockout-averse consumers (who repurchase at higher levels of inventory). The tendency to overestimate low inventory levels makes overstocking-averse consumers too conservative and thus limits their annual purchases to 504 units in the simulation, which is below their annual demand of 539 units. In contrast, the tendency to underestimate high inventory levels makes stockout-averse consumers too liberal and thus increases their annual purchases to 588 units, which is above their annual demand of 539 units. Which of these two segments should be targeted with sales promotions? To answer this question, we simulated the impact of a price reduction that would lower the purchase threshold by two units for both segments. We found that this would increase purchases by 12% among overstocking-averse consumers but by only 6% among stockout-averse consumers. The actual increase is likely to be even greater because these figures do not account for the consumption acceleration that higher inventory levels create. Because these overstocking-averse consumers probably include those with small households or those with small kitchens, multiunit promotions ("buy 5 get 1 free") and packaging would seem particularly appropriate because they facilitate flexible storage and reduce spoilage. Although identifying who the overstocking-averse consumers are would require specific data collection, the same reasoning would apply to category differences, suggesting benefits of targeting sales promotions to categories in which most consumers are averse to overstocking rather than to categories in which most consumers are averse to stockouts.

Finally, the simulation results show that retailers and manufacturers would increase their sales if they helped overstocking-averse consumers improve the accuracy of their estimations. This is because it would prevent these consumers from overestimating their low inventory levels and from mistakenly postponing their purchases. In contrast, improving the accuracy of stockout-averse consumers would reduce their purchases because it would prevent them from making the mistake of repurchasing when they still have a lot of product left in inventory.

What could be done to improve the accuracy of inventory estimations? The lack of self-knowledge about estimation strategies revealed by the protocol data and the robustness of the biases exhibited in the field studies suggest that consumers do not learn much from experience, even though they often run out of stock and waste overstocked products. (This, in itself, raises the issue as to whether consumers may be softening the negative consequences of estimation errors by adapting their consumption.)

Our model and our findings suggest that inventory estimations could be improved by helping consumers recognize when their estimations are likely to be particularly inaccurate, such as when their inventory is lower or higher than usual. In these cases, our findings suggest that consumers should increase the extremity of their estimations to compensate for the inelasticity of their intuitive estimates. Marketers could also improve the accuracy of inventory estimations by raising the perceptual salience of inventory levels. This could be done by changing where and how a product is stored, by usage-related advertising, or by transparent package designs that facilitate the monitoring of inventory.

REFERENCES

Abraham, Magid M. and Leonard M. Lodish (1993), "An Implemented System for Improving Promotion Productivity Using Store Scanner Data," *Marketing Science*, 12 (3), 248–69.

Ailawadi, Kusum and Scott A. Neslin (1998), "The Effect of Promotion on Consumption: Buying More and Consuming It Faster," *Journal of Marketing Research*, 35 (August), 390–98.

BBC News (2003), "William's Praise for St. Andrews," (May 30).

Bell, David R., Jeongwen Chiang, and V. Padmanabhan (1999), "The Decomposition of Promotional Response: An Empirical Generalization," *Marketing Science*, 18 (4), 504–526.

Chandon, Pierre and Brian Wansink (2002), "When Are Stockpiled Products Consumed Faster? A Convenience–Salience Framework of Postpurchase Consumption Incidence and Quantity," *Journal of Marketing Research*, 39 (August), 321–35.

——— and ——— (in press), "Is Obesity Caused by Calorie Underestimation? A Psychophysical Model of Fast-Food Meal Size Estimation," *Journal of Marketing Research*, forthcoming

Epley, Nicholas and Thomas Gilovich (2001), "Putting Adjustment Back in the Anchoring and Adjustment Heuristic," *Psychological Science*, 12 (5), 391–96.

——— and ——— (2004), "Are Adjustments Insufficient?" *Personality and Social Psychology Bulletin*, 30 (4), 447–60.

Frayman, Bruce J. and William E. Dawson (1981), "The Effect of Object Shape and Mode of Presentation on Judgments of Apparent Volume," *Perception and Psychophysics*, 29 (1), 56–62.

Green, David M. and John A. Swets (1988), *Signal Detection Theory and Psychophysics*. New York: Peninsula.

Jones, Timothy, Andrew Bockhorst, Brian McKee, and Aida Ndiaye (2003), "Percentage of Food Loss in the Household," report to the United States Department of Agriculture, Economics Research Service from the Bureau of Applied Research in Anthropology, University of Arizona.

Krider, Robert E., Priya Raghubir, and Aradhna Krishna (2001), "Pizzas: Pi or Square? Psychophysical Biases in Area Comparisons," *Marketing Science*, 20 (4), 405–425.

Krishna, Aradhna and Priya Raghubir (1997), "The Effect of Line Configuration on Perceived Numerosity of Dotted Lines," *Memory & Cognition*, 25 (4), 492–507.

Krueger, Lester E. (1982), "Single Judgments of Numerosity," *Perception & Psychophysics*, 31 (2), 175–82.

Lee, Eunkyu, Michael Y. Hu, and Rex S. Toh (2000), "Are Consumer Survey Results Distorted? Systematic Impact of Behavioral Frequency and Duration on Survey Response Errors," *Journal of Marketing Research*, 37 (February), 125–33.

Manning, Jeff (1999), *Got Milk? The Book*. Westminster, MD: Random House.

Menon, Geeta, Priya Raghubir, and Norbert Schwarz (1995), "Behavioral Frequency Judgments: An Accessibility-Diagnosticity Framework," *Journal of Consumer Research*, 22 (2), 212–28.

Meyer, Robert J. and Joao Assunçao (1990), "The Optimality of Consumer Stockpiling Strategies," *Marketing Science*, 9 (1), 18–41.

Mussweiler, Thomas, Fritz Strack, and Tim Pfeiffer (2000), "Overcoming the Inevitable Anchoring Effect: Considering the Opposite Compensates for Selective Accessibility," *Personality and Social Psychology Bulletin*, 26 (9), 1142–50.

Narasimhan, Chakravarthi, Scott A. Neslin, and Subrata Sen (1996), "Promotional Elasticities and Category Characteristics," *Journal of Marketing*, 60 (April), 17–30.

Pelham, Brett W., Tin Tin Sumarta, and Laura Myaskovsky (1994), "The Easy Path from Many to Much: The Numerosity Heuristic," *Cognitive Psychology*, 26 (April), 103–133.

P-O-P Times (1991), "Latest IRI Data Confirm the Effectiveness of P-O-P," (July–August), 37–42.

Raghubir, Priya and Eric A. Greenleaf (2006), "Ratios in Proportion: What Should the Shape of the Package Be?" *Journal of Marketing*, 70 (April), 95–107.

——— and Aradhna Krishna (1999), "Vital Dimensions in Volume Perception: Can the Eye Fool the Stomach?" *Journal of Marketing Research*, 36 (August), 313–26.

Roy, Michael M., Nicholas J.S. Christenfeld, and Craig R.M. McKenzie (2005), "Underestimating the Duration of Future Events: Memory Incorrectly Used or Memory Bias?" *Psychological Bulletin*, 131 (5), 738–56.

Stevens, Stanley Smith (1986), *Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects*. Oxford: Transaction Books.

Sun, Baohong (2005), "The Promotion Effect on Endogenous Consumption," *Marketing Science*, 24 (3), 430–43.

Teghtsoonian, Martha (1965), "The Judgment of Size," *American Journal of Psychology*, 78 (3), 392–402.

Wansink, Brian (2006), *Mindless Eating: Why We Eat More Than We Think*. New York: Bantam-Dell.

——— and Pierre Chandon (2006), "Can 'Low Fat' Nutrition Labels Lead to Obesity?" *Journal of Marketing Research*, 43 (November), forthcoming.

——— and Rohit Deshpandé (1994), "Out of Sight, Out of Mind: Pantry Stockpiling and Brand-Usage Frequency," *Marketing Letters*, 5 (1), 91–100.

———, Robert J. Kent, and Stephen J. Hoch (1998), "An Anchoring and Adjustment Model of Purchase Quantity Decisions," *Journal of Marketing Research*, 35 (February), 71–81.

——— and Koert Van Ittersum (2003), "Bottoms Up! The Influence of Elongation on Pouring and Consumption Volume," *Journal of Consumer Research*, 30 (3), 455–63.

Robert W. Palmatier, Rajiv P. Dant, Dhruv Grewal, & Kenneth R. Evans

Factors Influencing the Effectiveness of Relationship Marketing: A Meta-Analysis

Relationship marketing (RM) has emerged as one of the dominant mantras in business strategy circles, though RM investigations often yield mixed results. To help managers and researchers improve the effectiveness of their efforts, the authors synthesize RM empirical research in a meta-analytic framework. Although the fundamental premise that RM positively affects performance is well supported, many of the authors' findings have significant implications for research and practice. Relationship investment has a large, direct effect on seller objective performance, which implies that additional meditated pathways may explain the impact of RM on performance. Objective performance is influenced most by relationship quality (a composite measure of relationship strength) and least by commitment. The results also suggest that RM is more effective when relationships are more critical to customers (e.g., service offerings, channel exchanges, business markets) and when relationships are built with an individual person rather than a selling firm (which partially explains the mixed effects between RM and performance reported in previous studies).

Relationship marketing (RM), both in business practice and as a focus of academic research, has "experienced explosive growth" in the past decade (Srinivasan and Moorman 2005). Morgan and Hunt (1994, p. 22) define RM as "all marketing activities directed towards establishing, developing, and maintaining successful relational exchanges." Most research and practice assumes that RM efforts generate stronger customer relationships that enhance seller performance outcomes, including sales growth, share, and profits (Crosby, Evans, and Cowles 1990; Morgan and Hunt 1994), but some business executives have been disappointed in the effectiveness of their RM efforts (Colgate and Danaher 2000). Researchers have also suggested that in certain situations, RM may have a negative impact on performance (De Wulf, Odekerken-Schröder, and Iacobucci 2001; Hibbard et al. 2001).

Overall, these findings indicate that the effectiveness of RM efforts may vary depending on the specific RM strategy and exchange context; this inconsistency with regard to performance suggests the need for a meta-analysis to integrate the abundance of accumulated empirical research and to understand better the RM strategies that are most effective for building strong relationships, the outcomes that are most affected by customer relationships, and the conditions in which RM is most effective for generating positive seller outcomes. Advancing understanding of the primary drivers of RM effectiveness can increase the return on firms' RM investments dramatically and provide researchers with insights into ways to build more comprehensive models of the influence of RM on performance (Reinartz and Kumar 2003).

Using Dwyer, Schurr, and Oh's (1987) seminal article on relationships; Crosby, Evans, and Cowles's (1990) introduction of relationship quality; and Morgan and Hunt's (1994) key mediating variable theory of RM, most research has conceptualized the effects of RM on outcomes as fully mediated by one or more of the relational constructs of trust, commitment, relationship satisfaction, and/or relationship quality. The existing literature offers a wide range of antecedents for these relational mediators, and researchers disagree about which one best captures the characteristics of a relational exchange that influence performance. For example, Morgan and Hunt (1994) propose that trust and commitment are both key to predicting exchange performance, whereas others suggest that either trust (e.g., Doney and Cannon 1997; Sirdeshmukh, Singh, and Sabol 2002) or commitment (e.g., Anderson and Weitz 1992; Gruen, Summers, and Acito 2000; Jap and Ganesan 2000) alone is the critical relational construct.

Another school of thought suggests that the global construct of relationship quality, as reflected by a combination of commitment, trust, and relationship satisfaction, offers the best assessment of relationship strength and provides the most insight into exchange performance (e.g., De Wulf,

---

Robert W. Palmatier is an assistant professor, College of Business, University of Cincinnati (e-mail: rob.palmatier@uc.edu). Rajiv P. Dant is Frank Harvey Distinguished Professor of Marketing, College of Business Administration, University of South Florida (e-mail: rdant@coba.usf.edu). Dhruv Grewal is Toyota Chair of Commerce and Electronic Business and Professor of Marketing, Babson College (e-mail: dgrewal@babson.edu). Kenneth R. Evans is Professor of Marketing and Associate Dean, College of Business, University of Missouri (e-mail: evans@missouri.edu). The authors thank the Marketing Science Institute for its financial support of this research. They also thank Steven P. Brown, Elaine Allen, and Kent B. Monroe for their assistance on this project, as well as the anonymous *JM* reviewers for their constructive comments.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Odekerken-Schröder, and Iacobucci 2001; Kumar, Scheer, and Steenkamp 1995). These different relational mediators have been linked empirically to many antecedents and outcomes, which leads to the critical question, How does the relational mediated model vary across different relational perspectives?

In this article, we systematically review and analyze the literature on relational mediators in a meta-analytic framework (Figure 1) to provide insight into the following four research questions: (1) Which RM strategies are most effective for building customer relationships? (2) What outcomes are most affected by customer relationships? (3) Which moderators are most effective in influencing relationship–outcome linkages? and (4) How does the RM strategy → mediator → outcome linkage vary across different mediators?

Conceptual Framework

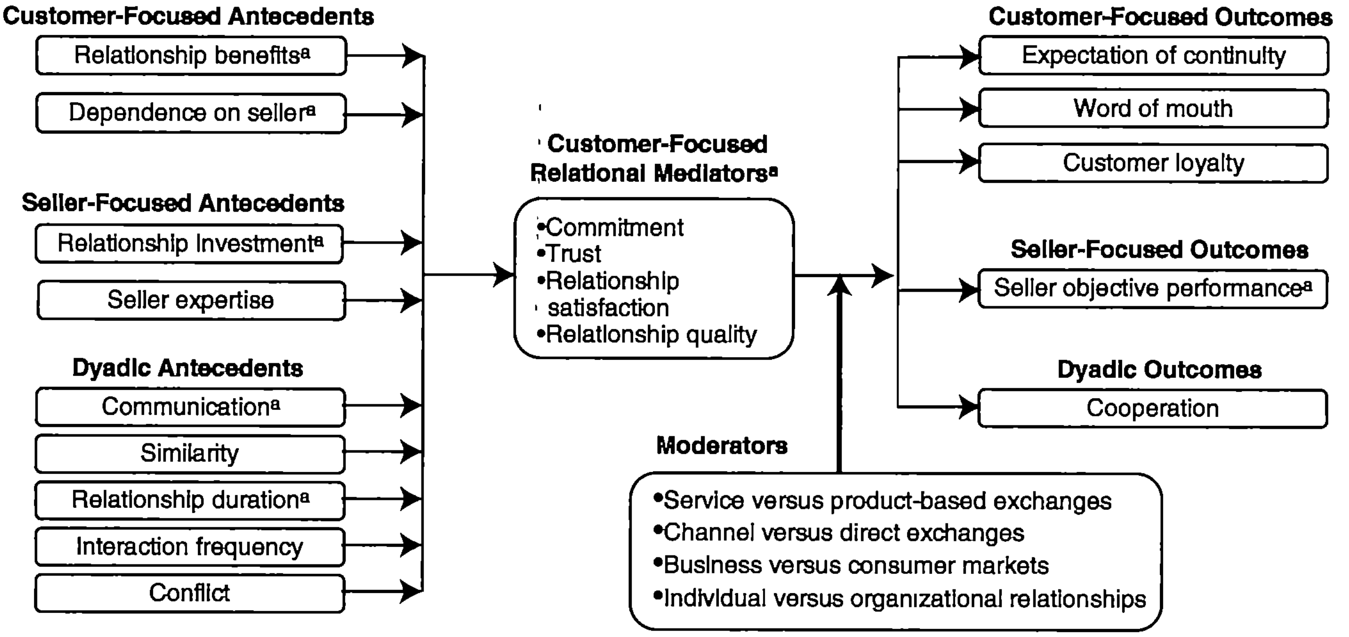
In reviewing the literature pertaining to relational mediators, we identified many constructs with similar definitions that operate under different aliases and constructs with similar names but different operationalizations. Thus, we use a single construct definition (see Table 1) to code existing research; we include a construct in the conceptual framework only if at least 10 effects emerge to support its empirical analysis. Of the many constructs investigated, only 18 met these criteria and appear in the model. Our nomological placement of each construct is driven by both theory and the frequency of placement in extant research. Of the studies that include hypothesized relationships with relational mediators, more than 90% are consistent with the causal ordering of constructs in our framework, with the exceptions of conflict and cooperation, which agree with our nomological framework in approximately 70% of extant studies.

Although a relationship is, by its very nature, two sided and both parties typically share in the benefits of a strong relationship, some antecedents and outcomes may have differential effects according to the measurement perspective (e.g., dependence). Thus, we adopt terminology to identify the perspective of each construct relative to its relational mediators. In this framework, "seller" refers to the party that implements the RM effort in the hope of strengthening its relationship with the "customer," and the relational mediator captures the customer's perception of its relationship with the seller. For clarity and consistency, we use these customer and seller perspectives even when the two parties may not be engaged in a typical exchange transaction (e.g., a strategic alliance). Thus, we classify antecedents and outcomes as "customer focused" when they share the same perspective as the relational mediator and as "seller focused" when they adopt a perspective opposite that of the evaluation of the relational mediator. We develop our conceptual framework in four parts, which roughly parallel our research questions, by first reviewing the literature on relational mediators, then investigating the antecedents and outcomes of these mediators, and, finally, studying potential moderators of the impact of relational mediators on outcomes.

Relational Mediators

Successful RM efforts improve customer loyalty and firm performance through stronger relational bonds (e.g., De Wulf, Odekerken-Schröder, and Iacobucci 2001; Sirdesh-

FIGURE 1
Relational Mediator Meta-Analytic Framework


Customer-Focused Antecedents
Relationship benefits[a]
Dependence on seller[a]
Seller-Focused Antecedents
Relationship investment[a]
Seller expertise
Dyadic Antecedents
Communication[a]
Similarity
Relationship duration[a]
Interaction frequency
Conflict
Customer-Focused Relational Mediators[a]
•Commitment
•Trust
•Relationship satisfaction
•Relationship quality
Moderators
•Service versus product-based exchanges
•Channel versus direct exchanges
•Business versus consumer markets
•Individual versus organizational relationships
Customer-Focused Outcomes
Expectation of continuity
Word of mouth
Customer loyalty
Seller-Focused Outcomes
Seller objective performance[a]
Dyadic Outcomes
Cooperation

[a] Construct had sufficient reported effects to be included in the multivariate causal model.

TABLE 1
Review of Construct Definitions, Aliases, and Representative Studies

Constructs	Definitions	Common Aliases	Representative Papers
Relational Mediators			
Commitment	An enduring desire to maintain a valued relationship	Affective, behavioral, obligation, and normative commitment	Anderson and Weitz 1992; Jap and Ganesan 2000; Moorman, Zaltman, and Deshpandé 1992; Morgan and Hunt 1994
Trust	Confidence in an exchange partner's reliability and integrity	Trustworthiness, credibility, benevolence, and honesty	Doney and Cannon 1997; Hibbard et al. 2001; Sirdeshmukh, Singh, and Sabol 2002
Relationship satisfaction	Customer's affective or emotional state toward a relationship, typically evaluated cumulatively over the history of the exchange	Satisfaction with the relationship, but not overall satisfaction	Crosby, Evans, and Cowles 1990; Reynolds and Beatty 1999
Relationship quality	Overall assessment of the strength of a relationship, conceptualized as a composite or multidimensional construct capturing the different but related facets of a relationship	Relationship closeness and strength	Crosby, Evans, and Cowles 1990; De Wulf, Odekerken-Schröder, and Iacobucci 2001
Antecedents			
Relationship benefits	Benefits received, including time saving, convenience, companionship, and improved decision making	Functional and social benefits and rewards	Hennig-Thurau, Gwinner, and Gremler 2002; Morgan and Hunt 1994; Reynolds and Beatty 1999
Dependence on seller	Customer's evaluation of the value of seller-provided resources for which few alternatives are available from other sellers	Relative and asymmetric dependence, switching cost, and imbalance of power	Hibbard, Kumar, and Stern 2001; Morgan and Hunt 1994
Relationship investment	Seller's investment of time, effort, spending, and resources focused on building a stronger relationship	Support, gifts, resources, investments, and loyalty programs	De Wulf, Odekerken-Schröder, and Iacobucci 2001; Ganesan 1994
Seller expertise	Knowledge, experience, and overall competency of seller	Competence, skill, knowledge, and ability	Crosby, Evans, and Cowles 1990; Lagace, Dahlstrom, and Gassenheimer 1991
Communication	Amount, frequency, and quality of information shared between exchange partners	Bilateral or collaborative communication, information exchange, and sharing	Anderson and Weitz 1992; Mohr, Fisher, and Nevin 1996; Morgan and Hunt 1994
Similarity	Commonality in appearance, lifestyle, and status between individual boundary spanners or similar cultures, values, and goals between buying and selling organizations	Salesperson or cultural similarity, shared values, and compatibility	Crosby, Evans, and Cowles 1990; Doney and Cannon 1997; Morgan and Hunt 1994
Relationship duration	Length of time that the relationship between the exchange partners has existed	Relationship age or length, continuity, and duration with firm or salesperson	Anderson and Weitz 1989; Doney and Cannon 1997; Kumar, Scheer, and Steenkamp 1995
Interaction frequency	Number of interactions or number of interactions per unit of time between exchange partners	Frequency of business contact and interaction intensity	Crosby, Evans, and Cowles 1990; Doney and Cannon 1997
Conflict	Overall level of disagreement between exchange partners	Manifest and perceived conflict or level of conflict, but not functional conflict	Anderson and Weitz 1992; Kumar, Scheer, and Steenkamp 1995

TABLE 1
Continued

Constructs	Definitions	Common Aliases	Representative Papers
Outcomes			
Expectation of continuity	Customer's intention to maintain the relationship in the future, which captures the likelihood of continued purchases from the seller	Purchase intentions, likelihood to leave (reverse), and relationship continuity	Crosby, Evans, and Cowles 1990; Doney and Cannon 1997
Word of mouth	Likelihood of a customer positively referring the seller to another potential customer	Referrals and customer referrals	Hennig-Thurau, Gwinner, and Gremler 2002; Reynolds and Beatty 1999
Customer loyalty	Composite or multidimensional construct combining different groupings of intentions, attitudes, and seller performance indicators	Behavioral loyalty and loyalty	De Wulf, Odekerken-Schröder, and Iacobucci 2001; Hennig-Thurau, Gwinner, and Gremler 2002; Sirdeshmukh, Singh, and Sabol 2002
Seller objective performance	Actual seller performance enhancements including sales, share of wallet, profit performance, and other measurable changes to the seller's business	Sales, share, sales effectiveness, profit, and sales performance	Reynolds and Beatty 1999; Sîguaw, Simpson, and Baker 1998
Cooperation	Coordinated and complementary actions between exchange partners to achieve mutual goals	Coordination and joint actions	Anderson and Narus 1990; Morgan and Hunt 1994

mukh, Singh, and Sabol 2002), but the literature offers varied perspectives on which relational constructs mediate the effects of RM efforts on outcomes. Commitment and trust are most often studied; commitment is "an enduring desire to maintain a valued relationship" (Moorman, Zaltman, and Deshpandé 1992, p. 316), and trust is "confidence in an exchange partner's reliability and integrity" (Morgan and Hunt 1994, p. 23). Another relationship mediator, relationship satisfaction, is a customer's affective or emotional state toward a relationship. Relationship satisfaction reflects exclusively the customer's satisfaction with the relationship and differs from the customer's satisfaction with the overall exchange. Other researchers have suggested that these mediators are merely indicators of the global mediator relationship quality, which is "an overall assessment of the strength of a relationship" and is conceptualized as a multidimensional construct that captures the many different facets of an exchange relationship (De Wulf, Odekerken-Schröder, and Iacobucci 2001, p. 36; see also Crosby, Evans, and Cowles 1990). Its structure and underlying dimensions vary across empirical studies, but central to the conceptualization is the belief that no single dimension or relational construct can fully define the "overall depth or climate" of an exchange relationship (Johnson 1999, p. 6).

Thus, whereas the literature consistently conceptualizes a mediating model for the effects of RM on performance, the specific relational mediators or composite of mediators appear to be driven mainly by researcher discretion; empirical comparisons of the differential effects of these relational mediators are noticeably absent. For example, some researchers propose trust as the critical relational mediator. Berry (1996, p. 42) offers "trust as perhaps the single most powerful relationship marketing tool available to a company," and Spekman (1988, p. 79) suggests that trust is the "cornerstone" of long-term relationships. Alternatively, Gundlach, Achrol, and Mentzer (1995, p. 78) propose commitment as the "essential ingredient for successful long-term relationships," and Morgan and Hunt (1994, p. 23) suggest "commitment among exchange partners as key to achieving valuable outcomes." De Wulf, Odekerken-Schröder, and Iacobucci (2001) prefer the overall concept of relationship quality to any specific component. In summary, there is little agreement among researchers as to which individual or composite relational mediator best captures the key aspects of a relationship that most affect outcomes. To address this issue empirically, our meta-analytic framework compares the relative effects of the different perspectives by analyzing relational mediators separately and as a group.

Antecedents to Relational Mediators

Customer-focused antecedents. Customers may perceive value in a relationship when they receive relationship benefits from an exchange partner (e.g., time savings, convenience, companionship), which increases their willingness to develop relational bonds. Relationship benefits have been shown to affect relational mediators positively (Morgan and Hunt 1994; Reynolds and Beatty 1999). Dependence on the seller reflects the customer's evaluation of the value of seller-provided resources for which few alternatives are available (Hibbard, Kumar, and Stern 2001). The literature is mixed regarding the effect of a customer's dependence, or relative dependence (i.e., the customer's dependence reduced by the seller's dependence), on relational mediators. Researchers find empirical support for both positive and negative influences of relative dependence on relational mediators (Anderson and Weitz 1989; Morgan and Hunt 1994), which indicates that its impact may be contingent on the context.

Seller-focused antecedents. Researchers have investigated various RM strategies that sellers can employ to strengthen relationships. Relationship investment refers to the time, effort, and resources that sellers invest in building stronger relationships. Such investments often generate expectations of reciprocation that can help strengthen and maintain a relationship and, therefore, positively influence relational mediators (Anderson and Weitz 1989; Ganesan 1994). Seller expertise reflects the knowledge, experience, and overall competence of the seller. When customers interact with a competent seller, they receive increased value, their relationship becomes more important, and they invest more effort to strengthen and maintain it (Crosby, Evans, and Cowles 1990; Lagace, Dahlstrom, and Gassenheimer 1991).

Dyadic antecedents. Customer- and seller-focused antecedents are meaningful from one side of the exchange dyad, but other antecedents require the active involvement of both exchange partners and are equally meaningful from both perspectives. For example, communication, or the amount, frequency, and quality of information shared between exchange partners (Mohr, Fisher, and Nevin 1996), requires both parties to exchange information. Communication builds stronger relationships in an exchange by helping resolve disputes, align goals, and uncover new value-creating opportunities (Morgan and Hunt 1994). Similarity is the commonality in appearance, lifestyle, and status between individual boundary spanners or the similar cultures, values, and goals between organizations. Such similarities between people or organizations may provide cues that the exchange partner will help facilitate important goals and has been shown to affect relational mediators positively (Crosby, Evans, and Cowles 1990; Doney and Cannon 1997). Relationship duration is the length of time that the relationship between the exchange partners has existed, whereas interaction frequency refers to the number of interactions per unit of time between partners. Both provide trading partners with more behavioral information in varied contexts, which allows for better predictions that should increase each party's confidence in its partner's behavior (Anderson and Weitz 1989; Doney and Cannon 1997). Finally, conflict entails the overall level of disagreement between exchange partners; this is often termed "perceived" or "manifest" conflict. As conflict increases, the customer is less likely to have confidence in the long-term orientation of the seller or to invest in building or maintaining a relationship; thus, conflict should negatively influence the customer's trust in and commitment toward the seller (Anderson and Weitz 1992).

Consequences of Relational Mediators

Customer-focused outcomes. Increased customer loyalty is one of the most common outcomes expected from RM efforts, but loyalty has been defined and operationalized in many different ways. An expectation of continuity reflects the customer's intention to maintain the relationship in the future and captures the likelihood of continued purchases. However, researchers have criticized this measure of loyalty because customers with weak relational bonds and little loyalty may report high continuity expectations as a result of their perceptions of high switching costs or their lack of time to evaluate alternatives (Oliver 1999). Word of mouth (WOM) captures the likelihood that a customer will refer a seller positively to another potential customer and, therefore, indicates both attitudinal and behavioral dimensions of loyalty. Some studies operationalize customer loyalty as a composite or multidimensional construct that includes groupings of intentions, attitudes, and seller performance indicators. We note that relational mediators positively influence global measures of customer loyalty, just as they do its individual components (De Wulf, Odekerken-Schröder, and Iacobucci 2001; Sirdeshmukh, Singh, and Sabol 2002).

Seller-focused outcomes. Possibly the most important outcome of RM efforts is seller objective performance, which captures the seller's actual performance enhancements, including sales, profit, and share of wallet. Some researchers have found empirical support for the influence of relational mediators on seller objective outcomes (e.g., Doney and Cannon 1997; Siguaw, Simpson, and Baker 1998), but several other studies have failed to find any significant effects, which implies that the effect of RM on performance may be context dependent (Crosby, Evans, and Cowles 1990; Gruen, Summers, and Acito 2000).

Dyadic outcomes. Cooperation captures the level of coordinated and complementary actions between exchange partners in their efforts to achieve mutual goals. Cooperation promotes value creation beyond that which each party could achieve separately, but because one party often receives its portion of the value earlier, the other party must have enough trust in the relationship to wait for its future reciprocation. Researchers have shown that trust and commitment between exchange partners are critical for cooperation (Anderson and Narus 1990; Morgan and Hunt 1994).

Moderators of Relational Mediators' Influence on Outcomes

The RM model (RM strategies → relational mediator → outcomes) we conceptualize herein can be applied across many different contexts in which business strategies may have varying effects. Therefore, an objective of our meta-analysis is to identify and empirically test the influence of potential moderators on the linkages in the RM model.

Contexts influencing relationship importance. Relationship marketing is based on the premise that building strong relationships positively influences exchange outcomes, and researchers recognize that exchanges vary across a spectrum from transactional to relational (Anderson and Narus 1991). For exchanges in which relationships are more important, we expect that the relational mediators will have a greater impact on outcomes, whereas in highly transactional exchanges, the relationships between buyers and sellers may have little influence on outcomes. Extant literature identifies three situations in which relationships may be more important for the success of an exchange. First, in general, services are perceived as less tangible, less consistent, and more perishable, and customers and sellers are more involved in the production and consumption of services than they are for products (Zeithaml, Parasuraman, and Berry 1985). This close interaction between customers and sellers may make customer–seller relationships more critical for services, and the intangibility of the offering may make the benefits of trust more critical because evaluations often are ambiguous.

Second, channel researchers tend to distinguish between channel partner exchanges and direct seller–customer transactions. Exchanges between channel partners have higher levels of interdependence, require coordinated action, and rely on the prevention of opportunistic behavior (Anderson and Weitz 1989). Thus, coordination improvements and the reduction of opportunistic behaviors through strong relationships should be more important in a channel context, which should lead to a greater impact of relational mediators on performance than their impact in direct exchanges.

Third, Anderson and Narus (2004, p. 21) differentiate consumer and business markets on the basis of the importance of relationships, maintaining that a "firm's success in business markets depends directly on its working relationships." If a working relationship is more critical for a customer's success in business markets, relationships should have a greater impact on exchange outcomes in business than in consumer markets.

Individual versus organizational relationships. Customers may form a relationship with an individual boundary spanner in the selling organization and/or with the selling organization as a whole. This issue of individual versus organizational relationships also has significant managerial implications as firms continue to try to increase their service efficiencies through the use of technology (e.g., customer relationship management). Experimental research shows that when people evaluate another individual, they make stronger, quicker, and more confident judgments than when they evaluate a group; those judgments also are more strongly related to outcomes and behaviors (Hamilton and Sherman 1996). Accordingly, we expect that customers' judgments based on the relational characteristics of an individual boundary spanner (e.g., trust in the salesperson) will be stronger, more confident, and more strongly linked to outcomes than their judgments based on the relational characteristics of a selling firm (e.g., trust in the firm). Post hoc findings support this premise. Doney and Cannon (1997, p. 45) report that "the process by which trust develops appears to differ when the target is an organization … as opposed to an individual salesperson," and Iacobucci and Ostrom (1996, p. 69) find that "[i]ndividual-to-firm relationships [are] also typically short-term and less intense in comparison to individual-level dyads." Thus, the positive effect of relational mediators on outcomes will be greater when the relational mediator is targeted toward an individual member of the selling organization than when it is targeted toward the organization.

Method

Collection and Coding of Studies

The key impetus for RM research was Dwyer, Schurr, and Oh's (1987) seminal article, so we searched empirical research for the mediators of interest during the period 1987–2004. We employed various methods in our literature search, including (1) a search of the ABI/Informs, PsycINFO, and Business Source Premier databases for each relational mediator; (2) a search of the Social Sciences Citation Index, using the seminal articles for these constructs; (3) manual shelf searches of journals that contain research on relational mediators; and (4) e-mails sent to researchers in the domain asking for their published and unpublished works. Our search generated more than 100 published and unpublished studies, each of which we evaluated for measures of the relationships among antecedents, outcomes, and the four relational mediators. Because correlations were the most common metric included in these studies (>95%), we e-mailed authors to request the correlation matrices for any studies in which they were not provided. Two independent coders, who were not familiar with the study, used the definitions in Table 1 to code the studies, and any differences (overall agreement >95%) were resolved through discussion (Szymanski and Henard 2001). When a single study provided more than one effect size estimate for the same relationship, we calculated an average. In cases in which the multiple effect size estimates from the same study were independent, we included them as separate effect size estimates. This procedure prevents the bias that may occur as a result of multiple counts of dependent effect size estimates and enables us to code moderators that vary across subsets of a sample in a single study (e.g., Brown and Peterson 1993). Ultimately, we combined 637 correlations from 111 independent samples drawn from 94 different manuscripts to yield a combined N of 38,077 with which to calculate the pairwise effect size estimates.[1]

Univariate Analyses

We began our analysis by adjusting our basic input measure, correlations (r), for corrections due to measurement error (scale reliability differences); we report the correlation adjusted for reliability (Hunter and Schmidt 1990). We then adjusted for sampling error (sample size differences), and we report the sample-weighted reliability-adjusted r and its 95% confidence intervals (CIs).[2] We calculated the chi-square test (d.f. = 1) for association and then addressed the file-drawer problem by computing the classical file drawer N (Rosenthal 1979) and the Q statistic test of homogeneity

---

[1] A list of the articles used in our empirical meta-analysis is available on request.

[2] Before applying the sample weights, we first converted the reliability-adjusted r's to variance-stabilizing Fisher's z scores (Rosenthal 1994; Shadish and Haddock 1994). Following standard procedures (Shadish and Haddock 1994, p. 268), we reconverted them back to r's to report the sample-weighted reliability-adjusted r and the 95% CIs.

(Cheung and Chan 2004; Hunter and Schmidt 1990) for each relationship.[3]

We performed such analyses for the influence of each antecedent on the four relational mediators (provided that there were four effects for each antecedent), which enables us to compare the influence of each antecedent on each mediator. Although these mediators measure different aspects of a relationship, researchers have argued that they are highly related and difficult to distinguish and, therefore, can be combined into a composite construct (Crosby, Evans, and Cowles 1990; De Wulf, Odekerken-Schröder, and Iacobucci 2001; Smith 1998). To facilitate the comparison of relative effects among the different antecedents on the overall relationship, we duplicated these analyses for the effects of each antecedent on all four mediators as a group.

Causal Model

In addition to the pairwise analyses, we aggregated the studies to test the nomological causal model implicit in Figure 1. This multivariate technique has the advantage of analyzing all linkages simultaneously, but it also needs significantly more data because the effects (i.e., correlation coefficients) must be available between each construct in the model and all other constructs, not just the pairwise effects for constructs with proposed relationships (Brown and Peterson 1993). Thus, causal models typically are limited to only the most commonly studied constructs. We determined the average-adjusted intercorrelation among all constructs in the framework whose required correlation coefficients were reported in three or more studies (Table 2). Furthermore, to increase the number of constructs that met this requirement and to provide a concise synthesis of the literature, we grouped all relational mediators together, and thus we propose a fully mediated model (Morgan and Hunt 1994). Of the 14 antecedent and outcome constructs we include in Figure 1, only 6 met this criterion and could be evaluated in the causal model.

Results

After we report the results of our causal model estimation procedure (Cheung and Chan 2005; Furlow and Beretvas 2005), we provide the results of the pairwise and casual model analyses structured around our four focal research questions. Because the first two questions focus on the effectiveness of antecedents that influence relational mediators (Table 3) and relational mediators that influence outcomes (Table 4), we report the findings beginning with the most influential constructs, and we concentrate on the aggregate results for all mediators (i.e., last row of each construct in Tables 3 and 4). Next, we report the results of the moderator analyses to understand the context in which RM is most effective. Finally, we concentrate on the last research question, namely, how the RM mediating model varies across different mediators. For this question, we no longer focus on the aggregate results but instead evaluate the effects that pertain to each mediator and thus provide insight into how the effects vary across mediators.

Causal Model Estimation

The fit indexes from the structural model testing of the causal model indicate that this model fits the data poorly: $\chi^2(5) = 322.27$, $p < .01$; comparative fit index = .87; goodness-of-fit index = .95; and root mean square error of approximation = .19 (for an example of this technique, see Brown and Peterson 1993). Modification indexes suggest a revised causal model that includes direct paths from dependence on seller and relationship investment to seller objective performance. The revised model results in indexes that indicate a good fit to the data: $\chi^2(3) = 12.15$, $p < .01$; comparative fit index = .89; goodness-of-fit index = .99; and root mean square error of approximation = .04.[4]

Which RM Strategies Are Most Effective for Building Customer Relationships?

As we show in Table 3, not all RM strategies (antecedents) are equally effective for building relationships. The average of the sample-weighted reliability-adjusted correlations among antecedents and mediators is .41, and they range from .13 for relationship duration to the largest absolute effect of –.67 for conflict. All paths from antecedents to relational mediators are supported in the pairwise analyses, except for the path from interaction frequency to relationship satisfaction. Most of these findings appear to be robust with regard to the number of null studies needed to render the observed effects zero (mean file-drawer N is 3152). Only three linkages appear susceptible to a file-drawer problem: interaction frequency → relationship satisfaction (N = 2), relationship duration → relationship satisfaction (N = 18), and relationship duration → relationship quality (N = 25). In the Q-statistic test for homogeneity, with one exception (i.e., seller expertise → relationship satisfaction), all the tests for homogeneity are significant.

Several insights can be drawn from the evaluation of the relative impact of different RM strategies on building strong

---

[3] To estimate the publication bias associated with published studies, we employed multiple methods: (1) Rosenthal's (1979) well-known file drawer method; (2) Orwin's (1983) failsafe N, which represents the number of missing studies (set to .05) that would bring the effect to .075, or less than the .10 effect level that Cohen (1977) classifies as signifying a low effect; and (3) comprehensive meta-analysis software (http://www.meta-analysis.com/html/stat_analysis_overview.html) to compute funnel plots for the various relationships. Funnel plots offer a simple scatter plot–based visual tool for investigating publication bias in meta-analyses (Sterne and Egger 2001). Overall, the funnel plots corroborate the inferences we drew from the file drawer N and Orwin's failsafe N; namely, the data we use in the meta-analyses do not display any evidence of publication bias.

[4] The median sample size from our meta-analysis of the studies included in the causal model is 2839. Modification indexes indicate that a direct path from relationship duration to seller objective performance could improve the model fit, but our evaluation of the parsimony-adjusted fit indexes suggests that the slight improvement in fit is more than offset by a loss in parsimony. The significance and pattern of effects do not change with this additional path, so we do not add it.

TABLE 2
Average Reliability-Adjusted Intercorrelations Among Constructs in Causal Model

Constructs	RBEN	DEPS	RINV	COMM	RDUR	RMED	SOP
Relation Benefits (RBEN)	[.87]						
SD							
Number of studies							
Cumulative sample size							
Dependence on Seller (DEPS)	.12	[.85]					
SD	.29						
Number of studies	4						
Cumulative sample size	886						
Relationship Investment (RINV)	.42	.13	[.82]				
SD	.08	.21					
Number of studies	7	5					
Cumulative sample size	1911	1273					
Communication (COMM)	.49	.28	.47	[.85]			
SD	.15	.21	.13				
Number of studies	10	12	9				
Cumulative sample size	2380	3260	2893				
Relationship Duration (RDUR)	.15	.14	.02	.18	[.99]		
SD	.03	.17	.06	.13			
Number of studies	3	7	7	6			
Cumulative sample size	1097	2150	3496	2282			
Relational Mediator (RMED)	.43	.19	.45	.51	.11	[.85]	
SD	.18	.25	.18	.19	.13		
Number of studies	18	33	24	38	26		
Cumulative sample size	5108	9296	8564	9803	10,720		
Seller Objective Performance (SOP)	.23	.29	.44	.30	.12	.35	[.92]
SD	.04	.31	.31	.30	.12	.22	
Number of studies	3	8	7	12	8	47	
Cumulative sample size	600	2860	2813	3149	4293	16,469	

Notes: Entries on the diagonal in brackets are weighted mean Cronbach's alpha coefficients. We included constructs in the causal model when three or more correlation coefficients were available among that construct and all other constructs in the model.

customer relationships. Conflict (r = –.67) has the largest absolute impact on the relational mediators of all antecedents, in support of the importance of resolving problems and disagreements to prevent relationship-damaging conflicts (alternatively, the presence of conflict may seriously undermine the effect of other RM antecedents). That the largest effect is negative extends to the RM domain the finding that people pay more attention to negatives than to positives (Shiv, Edell, and Payne 1997); this warrants further investigation. The seller expertise (r = .62) and communication (r = .54) antecedents have the greatest positive influence on relational mediators. The great impact of seller expertise suggests the importance of training boundary spanners and the potential detriments of staffing call centers with inexperienced or unskilled employees. The influence of seller expertise also seems to apply across all four relational mediators, in support of Vargo and Lusch's (2004, p. 3) claim that "skills and knowledge are the fundamental unit of exchange," such that sellers' skills and knowledge are the most important value-creating attributes. Similarly, the large positive effect of communication on all mediators is consistent with its role in both uncovering value-creating opportunities and resolving problems.

Relationship investment (r = .46), similarity (r = .44), and relationship benefits (r = .42) are the next most influential RM strategies. The strong positive impact of the seller's relationship investments and customer relationship benefits indicates that managers should engage in proactive RM spending. The importance of similarities between buyers and sellers suggests that without common reference points, exchange partners may find it difficult to move the exchange from a purely economic or transactional basis to a relational basis. The last three antecedents—dependence on seller (r = .26), interaction frequency (r = .16), and relationship duration (r = .13)—have notably smaller effects on relational mediators.

The causal model analysis generates the same rank order of relative effects of the antecedents on relational mediators as the pairwise analyses, which increases our confidence in the univariate results. Communication (β = .29, $p < .01$), relationship investment (β = .23, $p < .01$), relationship benefit (β = .18, $p < .01$), and dependence on seller (β = .05, $p < .01$) all have significant, positive effects on relational mediators, but relationship duration (β = .02) fails to influence relational mediators significantly in the multivariate analysis.

TABLE 3
Results: Descriptive Statistics and Influence of Antecedents on Relational Mediators

Proposed Relationships	Number of Raw Effects	Total N	Simple Average r	Average r Adjusted for Reliability	Sample-Weighted Reliability-Adjusted Average r	χ^2 for Association (d.f. = 1)	95% CI		File Drawer N (Using Two-Tailed Test)	Q Statistic for Homogeneity Test (d.f.)
							Lower Bound	Upper Bound		
Relationship benefits → commitment	11	3162	.37	.45	.51	987.50	.48	.53	2670	501.29 (10)
Relationship benefits → trust	13	3633	.31	.34	.33	420.99	.30	.36	1433	66.61 (12)
Relationship benefits → relationship satisfaction	7	2057	.41	.46	.45	490.59	.43	.49	916	53.56 (6)
Relationship benefits → relationship quality	8	2091	.33	.39	.36	353.47	.35	.43	749	60.27 (7)
Relationship benefits → all mediators	39	10,943	.35	.40	.42					
Dependence on seller → commitment	16	4670	.33	.40	.37	700.04	.35	.40	3206	530.30 (15)
Dependence on seller → trust	26	5935	.14	.17	.21	261.01	.18	.23	1354	452.24 (25)
Dependence on seller → relationship satisfaction	3	1076	.13	.13	.27	—	—	—	—	—
Dependence on seller → relationship quality	2	1604	.07	.08	.08	—	—	—	—	—
Dependence on seller → all mediators	47	13,285	.20	.24	.26					
Relationship investment → commitment	15	6544	.36	.43	.34	821.86	.32	.36	4200	381.19 (14)
Relationship investment → trust	17	4601	.38	.45	.45	1053.73	.42	.47	4455	161.70 (16)
Relationship investment → relationship satisfaction	10	2691	.42	.47	.52	882.85	.49	.55	2217	276.14 (9)
Relationship investment → relationship quality	9	2635	.49	.57	.58	1125.80	.55	.60	2750	245.40 (8)
Relationship investments → all mediators	51	16,471	.40	.47	.46					
Seller expertise → commitment	1	177	.78	.90	.90	—	—	—	—	—
Seller expertise → trust	12	3464	.49	.59	.52	1121.22	.49	.54	3988	279.45 (11)
Seller expertise → relationship satisfaction	5	1049	.49	.56	.56	413.64	.52	.60	473	1.20 (4)
Seller expertise → relationship quality	1	1009	.90	.98	.98	—	—	—	—	—
Seller expertise → all mediators	19	5699	.53	.62	.62					
Communication → commitment	25	5840	.45	.53	.55	2244.96	.54	.57	13,712	323.37 (24)
Communication → trust	29	7948	.43	.51	.56	3146.74	.54	.57	21,962	939.32 (28)
Communication → relationship satisfaction	6	1727	.46	.51	.51	546.15	.48	.55	870	34.78 (5)
Communication → relationship quality	7	1907	.35	.40	.43	407.26	.40	.47	709	82.39 (6)
Communication → all mediators	67	17,422	.43	.51	.54					
Similarity → commitment	3	386	.54	.66	.63	—	—	—	—	—
Similarity → trust	10	2562	.50	.56	.41	475.20	.38	.44	1884	347.42 (9)
Similarity → relationship satisfaction	1	151	.30	.34	.34	—	—	—	—	—
Similarity → relationship quality	1	1009	.22	.24	.24	—	—	—	—	—
Similarity → all mediators	15	4108	.48	.54	.44					

TABLE 3
Continued

Proposed Relationships	Number of Raw Effects	Total N	Simple Average r	Average r Adjusted for Reliability	Sample-Weighted Reliability-Adjusted Average r	χ^2 for Association (d.f. = 1)	95% CI		File Drawer N (Using Two-Tailed Test)	Q Statistic for Homogeneity Test (d.f.)
							Lower Bound	Upper Bound		
Relationship duration → commitment	13	6638	.12	.13	.11	82.43	.09	.14	312	241.99 (12)
Relationship duration → trust	20	8201	.12	.12	.14	162.63	.12	.16	674	79.73 (19)
Relationship duration → relationship satisfaction	5	1542	.09	.09	.13	24.15	.08	.17	18	11.14 (4)
Relationship duration → relationship quality	5	1830	.11	.11	.11	22.95	.07	.16	25	25.15 (4)
Relationship duration → all mediators	43	18,211	.11	.12	.13					
Interaction frequency → commitment	2	724	–.03	–.03	–.03	—	—	—	—	—
Interaction frequency → trust	10	2198	.30	.33	.30	210.97	.26	.35	680	294.63 (9)
Interaction frequency → relationship satisfaction	4	965	.11	.11	.04	1.47	–.02	.10	2	27.81 (3)
Interaction frequency → relationship quality	3	1124	–.03	–.03	–.03	—	—	—	—	—
Interaction frequency → all mediators	19	5011	.17	.19	.16					
Conflict → commitment	10	4339	–.41	–.49	–.71	3331.03	–.72	–.69	5496	1145.07 (9)
Conflict → trust	9	2906	–.54	–.63	–.65	1769.52	–.68	–.63	4033	206.98 (8)
Conflict → relationship satisfaction	1	95	–.27	–.31	–.31	—	—	—	—	—
Conflict → relationship quality	0	—	—	—	—	—	—	—	—	—
Conflict → all mediators	20	7340	–.46	–.55	–.67					

Notes: Operationally, we attempted calculations only when there was a minimum of four raw effects associated with a relationship. A dash indicates that this condition was not met.

TABLE 4
Results: Descriptive Statistics and Influence of Relational Mediators on Outcomes

Proposed Relationships	Number of Raw Effects	Total N	Simple Average r	Average r Adjusted for Reliability	Sample-Weighted Reliability-Adjusted Average r	χ^2 for Association (d.f. = 1)	95% CI Lower Bound	95% CI Upper Bound	File Drawer N (Using Two-Tailed Test)	Q Statistic for Homogeneity Test (d.f.)
Commitment → expectation of continuity	16	4215	.45	.54	.53	1447.81	.51	.55	5895	210.88 (15)
Trust → expectation of continuity	24	6632	.47	.55	.58	2889.95	.56	.60	15,456	274.10 (23)
Relationship satisfaction → expectation of continuity	5	1879	.50	.58	.57	778.08	.54	.60	949	21.36 (4)
Relationship quality → expectation of continuity	3	1733	.50	.55	.54	—	—	—	—	—
All mediators → expectation of continuity	48	14,459	.47	.55	.56					
Commitment → WOM	6	3674	.52	.61	.64	2111.52	.62	.66	2707	147.41 (5)
Trust → WOM	5	3507	.48	.56	.62	1833.48	.60	.64	1804	61.05 (4)
Relationship satisfaction → WOM	3	1054	.48	.50	.53	—	—	—	—	—
Relationship quality → WOM	3	1733	.58	.61	.60	—	—	—	—	—
All mediators → WOM	17	9968	.51	.58	.61					
Commitment → customer loyalty	12	4588	.45	.54	.58	1996.01	.56	.60	5447	151.79 (11)
Trust → customer loyalty	20	6328	.44	.51	.54	2248.51	.52	.55	10,572	308.98 (19)
Relationship satisfaction → customer loyalty	9	2781	.35	.39	.41	522.61	.38	.44	1188	129.89 (8)
Relationship quality → customer loyalty	9	2851	.40	.46	.47	750.86	.45	.50	1722	55.04 (8)
All mediators → customer loyalty	50	16,548	.42	.48	.52					
Commitment → seller objective performance	20	7342	.30	.35	.27	549.90	.25	.29	3489	407.99 (19)
Trust → seller objective performance	32	10,306	.29	.33	.35	1333.28	.33	.36	10,108	924.64 (31)
Relationship satisfaction → seller objective performance	7	1605	.32	.37	.32	172.58	.27	.36	376	84.62 (6)
Relationship quality → seller objective performance	6	3517	.28	.31	.63	1930.18	.61	.65	1986	2,641.29 (5)
All mediators → seller objective performance	65	22,770	.29	.34	.35					
Commitment → cooperation	16	4436	.41	.50	.64	2509.07	.62	.66	7385	418.42 (15)
Trust → cooperation	24	6192	.56	.67	.73	5340.80	.72	.74	28,898	476.68 (23)
Relationship satisfaction → cooperation	5	931	.45	.55	.68	630.30	.64	.71	468	17.33 (4)
Relationship quality → cooperation	0	—	—	—	—	—	—	—	—	—
All mediators → cooperation	45	11,559	.50	.60	.70					

Notes: Operationally, we attempted calculations only when there was a minimum of four raw effects associated with a relationship. A dash indicates that this condition was not met.

What Outcomes Are Most Affected by Customer Relationships?

In Table 4, we show that customer relationships do not equally influence all exchange outcomes. The average of the correlations among relational mediators and outcomes is .55, ranging from a low of .35 for seller objective performance to a high of .70 for cooperation. All paths from relational mediators to outcomes are supported. None of these results appears to be susceptible to a file-drawer problem; all paths would require more than 375 null studies to generate a zero effect, with a mean file drawer N of 6153. All Q-statistic tests for homogeneity are significant, demonstrating statistical heterogeneity and supporting a moderator analysis.

Relational mediators have the largest combined influence on the dyadic outcome of cooperation (r = .70), followed by WOM (r = .61). This finding reinforces the importance of relationship building for a high level of customer cooperation. The greater impact of relational mediators on WOM (r = .61) than on the expectation of continuity (r = .56) or on customer loyalty (r = .52) lends support to Reichheld's (2003, p. 48) premise that WOM may be the best indicator of "intense loyalty." Only customers who have strong relationships with sellers are willing to risk their own reputation by giving a referral.

Of the five outcomes, relational mediators have the least influence on seller objective performance (r = .35). Thus, although customer relationships positively influence performance outcomes, in support of efforts put into RM strategies, the actual effect on performance is lower than that on the other four outcomes. This finding is not surprising; relational mediators are more closely related to loyalty and cooperation than is objective performance, which often depends on other, nonrelational factors (e.g., the economy).

The causal model includes only one outcome, but we confirm the significant influence of relational mediators on seller objective performance ($\beta = .16$, $p < .01$). In addition, although the impact of relationship benefits and communication strategies on seller objective performance is fully mediated by the relational mediators, the influence of dependence and relationship investment is only partially mediated; both dependence on seller ($\beta = .22$, $p < .01$) and relationship investment ($\beta = .34$, $p < .01$) also have large, direct effects on seller performance.

Which Moderators Are Most Effective in Influencing Relationship–Outcome Linkages?

In Table 5, we present the influence of moderators on the linkage between relational mediators and outcomes.[5] The premise that customer relationships have a greater impact on exchange outcomes in situations in which relationships are more critical to the success of the exchange is supported for the impact of all mediators on customer loyalty among services, channels, and business customers. The correlation of all mediators with customer loyalty is .58 for service versus .43 for product-based exchanges ($p < .05$), .65 for channel versus .46 for direct interactions ($p < .01$), and .56 for business versus .46 for consumer markets ($p < .05$). We find a similar effect in business markets for the impact of relationships on seller objective performance, for which the influence of all mediators is r = .36 in business markets versus r = .25 in consumer markets ($p < .01$). In summary, the significant moderation of the influence of relationships on customer loyalty across services, channels, and business markets, as well as on performance in business markets, provides support for our premise that customer relationships have a greater impact on exchange outcomes in situations in which relationships are more critical to success.

Contrary to our expectations, relational mediators' influence on the expectation of continuity is greater in consumer than in business markets, mostly because of commitment's influence on the expectation of continuity (r = .46 business, r = .71 consumer; $p < .05$). Because commitment taps a customer's desire to maintain a valued relationship, whereas the expectation of continuity captures a customer's intent to maintain the relationship, consumers may be better able to convert their attitudes or desire into an intention than are business buyers because consumers have a higher degree of control over their actions. Consistent with the theory of planned behavior, the link between an attitude and an intention should be stronger as control increases (Ajzen and Fishbein 1980). Thus, the stronger impact of commitment on the expectation of continuity (which results from higher levels of control) in consumer markets than in business markets may offset the typically greater importance of relationships in business markets.

As we proposed, relationships have a greater impact on customer loyalty when the target of the relationship is an individual person (r = .56) than when the target is an organization (r = .46; $p < .05$). Similarly, the impact of relational mediators on cooperation is greater (r = .68 for interpersonal, r = .55 for interorganizational; $p < .05$) when the customer's relationship is targeted toward a person employed by the seller than when it is targeted toward the seller overall. We provide additional support for this finding in Table 5, in which we show that of the 16 moderation tests, 81% are in the expected direction, and the impact of all mediators on seller objective performance is significant at the $p < .10$ level (r = .40 for interpersonal, r = .31 for interorganizational).

How Does the RM Strategy → Mediator → Outcome Linkage Vary Across Mediators?

The preceding research questions focus on the effects of the four mediators as a group. In this subsection, we investigate the individual linkages to identify when mediators operate differently; we begin with the front half of the model: the RM strategy → relational mediator linkage (Table 3). The effectiveness of RM strategies varies across different relational mediators. We consider the differential effects of relationship investments and benefits on mediators together because they are logically related. Sellers' relationship

---

[5] We carried out the moderator analysis using the procedure that Brown (1996) and Grewal and colleagues (1997) employ. However, the results we report in Table 5 must be interpreted cautiously because in the majority of the nonsignificant cases, the power of the test is relatively small (Cohen 1977; Fern and Monroe 1996). On the basis of our power analysis, we have flagged the relationships in Table 5 that we believe researchers would be premature in dismissing.

TABLE 5
Influence of Moderators on Relational Mediators' Effects on Outcomes

Moderated Relationships	Total Number of Raw Effects	Service Versus Product-Based Exchanges		Channel Versus Direct Exchanges		Business Versus Consumer Markets		Individual Versus Organizational Relationships	
		Service	Product	Channel	Direct	Business	Consumer	Individual	Organizational
Commitment → expectation of continuity	16	.63 (6)	.46 (8)[a]	.54 (4)	.53 (10)	.46 (11)	.71 (5)*	.54 (4)	.53 (12)
Trust → expectation of continuity	24	.51 (6)	.56 (15)	.57 (7)	.55 (13)	.54 (15)	.57 (9)	.50 (4)	.55 (17)
Relationship satisfaction → expectation of continuity	5	—	—	—	—	—	—	—	—
Relationship quality → expectation of continuity	3	—	—	—	—	—	—	—	—
All mediators → expectation of continuity	48	.56 (15)	.53 (24)	.54 (26)	.57 (12)	.52 (33)	.61 (15)*	.52 (13)	.55 (32)
Commitment → customer loyalty	12	.70 (2)	.49 (8)*	.63 (1)	.52 (9)[a]	.59 (3)	.52 (9)[a]	.62 (2)	.52 (10)*
Trust → customer loyalty	20	.53 (5)	.47 (12)	.68 (1)	.49 (17)[a]	.49 (4)	.51 (16)	.51 (5)	.50 (14)
Relationship satisfaction → customer loyalty	9	—	—	—	—	.60 (2)	.33 (7)*	.59 (2)	.33 (7)*
Relationship quality → customer loyalty	9	.54 (1)	.40 (6)[a]	—	—	.60 (2)	.42 (7)*	.61 (1)	.44 (8)*
All mediators → customer loyalty	50	.58 (8)	.43 (33)*	.65 (2)	.46 (40)*	.56 (11)	.46 (39)*	.56 (10)	.46 (39)*
Commitment → seller objective performance	20	26 (6)	.40 (10)[a]	.37 (7)	.37 (11)	.39 (16)	.21 (4)*	.28 (2)	.36 (18)
Trust → seller objective performance	32	.27 (4)	.34 (24)	.31 (7)	.30 (16)	.35 (26)	.26 (6)	.38 (13)	.31 (17)
Relationship satisfaction → seller objective performance	7	—	—	—	—	.40 (5)	.29 (2)	.44 (5)	.19 (2)[a]
Relationship quality → seller objective performance	6	—	—	—	—	—	—	.59 (2)	.17 (4)[a]
All mediators → seller objective performance	65	.32 (12)	.36 (41)	.36 (19)	.34 (31)	.36 (53)	.25 (12)*	.40 (22)	.31 (41)[a]
Commitment → cooperation	16	.63 (3)	.47 (11)[a]	.49 (7)	.51 (9)	.50 (15)	.54 (1)	.79 (2)	.46 (14)*
Trust → cooperation	24	.73 (3)	.67 (21)	.70 (10)	.66 (14)	.67 (23)	.66 (1)	.70 (9)	.64 (14)
Relationship satisfaction → cooperation	5	—	—	—	—	—	—	—	—
Relationship quality → cooperation	0	—	—	—	—	—	—	—	—
All mediators → cooperation	45	.67 (7)	.59 (36)[a]	.60 (21)	.60 (24)	.60 (42)	.61 (3)	.68 (14)	.55 (30)*

*$p < .05$ (one-tailed).

[a]Nonsignificant results should be interpreted cautiously. In a majority of the cases for which a moderator variable does not significantly moderate the effect of a given relationship, the power associated with the test is relatively small. On the basis of the expected effect size, power, and number of studies required to move the power to 80%, we identify (with superscript "a") relationships that researchers may be premature in dismissing as not significantly moderated (Cohen 1977, Fern and Monroe 1996).

Notes: The cell entries show the average effects encountered for each moderator level, with the total number of effects in parenthesis, subjected to t-test comparisons (Brown 1996). We dropped studies that could not be coded into subgroups from the comparisons. In addition, the limited number of effects suggested dropping WOM from the analysis. Operationally, we carry out comparisons only when the total number of raw effects is six or more to ensure the a priori probability of finding at least three effects at each level of moderator. When this cutoff is not met or when the number of effects for one level of moderator is less than 1, we use a dash to indicate that we did not perform that particular moderator analysis.

investments normally generate customer relationship benefits, but in some cases, an investment may not be desired or generate any actual benefit. Relationship investment has the least impact on commitment (r = .34) of all the relational mediators (mean of other relational mediators, r = .52), with no overlap in the CIs. Thus, sellers can strengthen their overall relationships through investments (possibly by generating feelings of reciprocity), but the relative impact on customer commitment is minimal. Alternatively, customer relationship benefits have the greatest impact on customer commitment (r = .51), especially compared with customer trust (r = .33, no overlap in CI), which suggests that customers value these benefits and want to maintain them. This discrepancy may occur because many relationship investments do not generate value for the customer and therefore do not lead to customer commitment. Although investments that do not generate customer value may strengthen relationships by generating debts of reciprocity, they will not necessarily generate an enduring desire to maintain a valued relationship.

As we might have expected, dependence has a greater positive effect on commitment (r = .37) than the other mediators (mean of other relational mediators, r = .19, no overlap in CI), which reflects customers' desire to maintain a relationship with the seller on which they are dependent. The relatively limited effect of dependence on customer trust (r = .21) may be due to customers' concerns that sellers will take advantage of their dependence.

Although similarity often is hypothesized to influence trust by reducing uncertainty and serving as a cue to facilitate goals, we find that similarity actually has a greater impact on commitment (r = .63) than on trust (r = .41). This greater impact on commitment might be explained by research on stereotype behaviors, which suggests that people want to strengthen and maintain relationships with "in-group" members and that similarity is a proxy for customers' perceptions of a seller's fit with their in-group (Devine 1995).

The influence of interaction frequency on trust is much greater (r = .30) than that of the three other mediators (mean of other relational mediators, r = –.01, no overlap in CI). As customers interact more frequently with sellers, they appear to gain more information about their partner, which reduces their uncertainty about future behaviors and improves trust; however, the frequency of their interaction has little effect on other relational mediators.

We now turn our attention to the back half of the model, the relational mediator → outcome linkage (Table 4), for which we find that relational mediators have differential effects on most of the outcomes studied. Commitment (r = .58) has the greatest influence on customer loyalty (mean of all other relational mediators, r = .47, no overlap of CI), as we might expect from these two similar constructs.

Moreover, relationship quality has the greatest influence on objective performance (r = .63), followed by trust (r = .35), relationship satisfaction (r = .32), and commitment (r = .27), and the CIs of relationship quality, trust, and commitment do not overlap. These findings indicate that RM researchers may need to take a multiple mediator or composite view when they measure customers' relationships to capture their impacts on objective performance. Different dimensions of a relationship may be synergistic, and superior performance may be possible only when the relationship is sufficiently strong on all critical aspects. Trust (r = .73) is most critical for cooperation compared with the other mediators (mean of other relational mediators, r = .66, no overlap of CI), in support of its role in coordinating actions among partners to create value and achieve mutual outcomes.

Discussion

We provide evidence that the intervening role of relational mediators between RM strategies and exchange outcomes is more complex than is currently suggested in the extant research, but the fundamental premise that RM and strong relationships positively affect performance is well supported. Several of our findings offer important implications for improving the effectiveness of RM research and practice (for a summary of key findings and implications, see Table 6).

First, RM strategies/antecedents have a wide range of effectiveness in terms of generating strong relationships, though specific strategies appear to be most effective for strengthening specific aspects of a relationship. Overall, expertise and communication are the most effective relationship-building strategies across all elements of a relationship, whereas the other strategies often have differential effects across the different mediators. For example, generating relationship benefits, promoting customer dependency, and increasing similarity to customers are more effective strategies for increasing customer commitment than for building trust, whereas relationship investment and interaction frequency have the opposite effect. Therefore, when comparing the relative effectiveness of RM strategies, the results depend on the relational mediator under investigation. These findings indicate that RM may be improved by taking a more fine-grained approach in which managers target RM strategies at specific relational weaknesses.

Second, we find that objective performance is influenced most by relationship quality (a composite measure of relationship strength) and least by commitment, which supports a multidimensional perspective of relationships in which no single or "best" relational mediator can capture the full essence or depth of a customer–seller relationship (Hennig-Thurau, Gwinner, and Gremler 2002; Johnson 1999). Previous research (Berry 1996; Doney and Cannon 1997; Spekman 1988) that offers either commitment or trust as the key, central, or cornerstone relational mediator may be focused too narrowly; a relationship may be truly effective only when most or all of its key aspects are strong. Therefore, research that focuses only on commitment and generalizes from its impact on customer intention or intermediate behaviors to its effect on seller performance may prove misleading. For example, commitment has the greatest impact on customer loyalty and the smallest impact on objective performance.

Third, the large, direct effects of dependence and relationship investment on seller objective performance in the causal model suggest that these antecedents influence performance through alternative, mediated pathways. Although

TABLE 6
Summary of Key Findings and Implications

Key Findings	Research and Managerial Implications
Antecedents	
Relationship marketing strategies/antecedents have a wide range of effectiveness for generating strong relationships. Expertise and communication are most effective, then relationship investment, similarity, and relationship benefits; dependence, frequency, and duration are relatively ineffective.	Selection and training of boundary spanners is critical; expertise, communication, and similarity to customers are some of the most effective relationship-building strategies. Expertise's impact supports Vargo and Lusch's (2004) premise that "skills and knowledge" are the most important seller value-creation attributes.
The negative impact of conflict is larger in magnitude than the positive effect of any other RM strategy.	All proactive RM efforts may be wasted if customer conflict is left unresolved.
Specific RM strategies appear most effective for strengthening one aspect of a relationship. Relationship benefits, customer dependency, and similarity are more effective for increasing commitment than for building trust; the opposite is true for relationship investment and frequency.	Relationship marketing may be improved through a fine-grained approach that targets specific relational weaknesses. The relative effectiveness of RM strategies depends on the relational mediator investigated.
Outcomes	
Relationship quality (a composite measure of relationship strength) has the greatest influence on objective performance, and commitment has the least.	No single relational mediator captures the full essence or depth of a customer–seller relationship; the findings support a multidimensional perspective of relationships. Extant research focused on a single relational mediator may provide misleading guidance.
Surprisingly, relationship investment has a large, direct effect on seller objective performance, in addition to its frequently hypothesized indirect mediated effect.	The classic mediating model of RM (Morgan and Hunt 1994) should be adapted to include alternative mediated pathways (e.g., reciprocity).
Dependence has a large, direct effect on seller objective performance but a relatively small impact on relational mediators.	Dependence is not an effective relationship-building strategy but can improve performance in other ways, possibly by increasing switching costs and barriers to exit.
Of all outcomes, relationships have the greatest influence on cooperation and WOM and the least on objective performance.	Relationship marketing efforts may be effectively extended across many other nontraditional buyer–seller interactions (e.g., interdepartmental groups) for which cooperation is often critical for success. Word-of-mouth behaviors may be the best discriminator of true customer loyalty (Reichheld 2003).
Moderators	
Relationship marketing is typically more effective when relationships are more critical to customers, such as for (1) service versus product offerings, (2) channel versus direct exchanges, and (3) business versus consumer markets.	Researchers must take care when extending findings across contexts in which relationship importance may vary. Managers might target RM expenditures to customer segments with the highest desire for strong relationships to improve returns.
Customer relationships often have stronger effects on exchange outcomes when their target is an individual person than when it is a selling firm.	Researchers should differentiate the effects of customer relationships with boundary spanners from those with firms. Strategies such as team selling, salesperson disintermediation, and the use of call centers should be evaluated in light of the impact of interpersonal relationships.

dependence is not very effective at building relationships, it can improve performance by increasing switching costs and barriers to exit, which may make it an effective performance-enhancing strategy but not an effective RM strategy. However, relationship investment both builds customer relationships and directly improves performance, which suggests that the extant relational-mediated framework is not comprehensive and that additional mediators (e.g., reciprocity) must be investigated to explain the impact of RM on performance fully. The importance of capturing the direct effect of relationship investment ($\beta = .34$) is reinforced by its greater impact on objective performance than the effect of the relational mediators ($\beta = .16$) in the causal model.

Fourth, the findings that strong relationships appear to be more effective for building customer loyalty and improving seller performance for (1) service versus product offerings, (2) channel versus direct exchanges, and (3) business versus consumer markets lend support to the premise that RM may be a more effective strategy in situations in which relationships are more critical. This finding calls into question sellers' efforts to force RM strategies in contexts in which the customer's relational needs are unclear; it also may explain the less-than-desirable results of RM on performance that have been documented in previous studies (e.g., Reinartz and Kumar 2003). Because these situational moderators are coarse proxies for customers' relationship needs, RM effectiveness likely varies across other factors that influence customers' needs for strong relationships. In turn, researchers must take care when extending RM research to these different contexts.

Fifth, the results suggest that customer relationships have stronger effects on exchange outcomes when their target is an individual person than when their target is a selling firm. Thus, RM strategies focused on building interpersonal relationships between boundary spanners (e.g., dedicated salesperson, social entertaining) may be more effective than those focused on building customer–firm relationships (e.g., team selling, frequency-driven loyalty programs). Social psychology's individual and group judgment theory (Hamilton and Sherman 1996, p. 336), which posits "differences in the outcomes of impressions formed of individual and group targets, even when those impressions are based on the very same behavioral information," has several implications for the marketing domain and may provide a parsimonious explanation for previous marketing research (Doney and Cannon 1997; Iacobucci and Ostrom 1996). The post hoc finding that conflict has a more negative impact ($p < .01$) on customer–firm relationships than on customer–individual relationships is also consistent with this theory because judgments about individuals are more resilient to disconfirming events than are judgments about groups (Hamilton and Sherman 1996). Thus, managers may want to use boundary spanners or salespeople rather then centralized service centers to resolve conflicts because customers' relationships with salespeople may withstand conflict better than their relationships with selling firms.

Managerial Implications

Most promising for managers is that five of the strategies with the greatest impact are either seller focused or dyadic, in support of the effectiveness of proactive relationship-building strategies undertaken by sellers. Business executives focused on building and maintaining strong customer relationships should note that the selection and training of boundary spanners is critical; expertise, communication, and similarity to customers are the most effective relationship-building strategies. The next most effective strategy is for managers to make relationship investments and generate relationship-based benefits for customers; furthermore, relationship investment has the added benefit of influencing performance directly. However, managers must recognize that these proactive efforts will be wasted if they leave customer conflict unresolved because the negative influence of conflict on customer relationships is greater in magnitude than that of any other strategy. Thus, some firms could generate higher returns by reallocating their RM investments to conflict resolution. Extending service recovery research into the RM domain to develop strategies for "relationship recovery" also might be worthwhile. A strategy of increasing customer dependence does not appear to be an effective way to build relationships, but it seems to influence seller performance directly. Neither relationship duration nor interaction frequency is a good driver of strong customer relationships.

Of all the outcomes we analyze, relationships have the greatest influence on cooperation and WOM. The impact on cooperation implies that RM efforts may be effectively extended across many nontraditional buyer–seller interactions (e.g., alliances, interdepartmental groups); in these situations, cooperation is often critical for success. Simply stated, firms that depend on WOM strategies for new customers should implement effective RM programs.

Some results indicate that a more targeted effort may improve RM efficiency. Because RM strategies appear to operate through different mediators that affect outcomes differentially, a manager who desires cooperation between two groups after a merger and who recognizes that trust is the relational mediator with the greatest influence on cooperation should select RM strategies that influence trust best (i.e., communication and interaction frequency). Marketers with a portfolio of customers, channels, and products could improve the return on their RM expenditures by targeting their spending toward segments in which RM is more likely to pay off, such as customers who purchase more services, channel versus direct customers, and business versus consumer segments.

Managers may also want to leverage the potentially stronger impact on customer loyalty and seller performance in relationships that involve an individual boundary spanner. For firms that experience low turnover, focusing their RM efforts on building customer–salesperson bonds may be a productive strategy, though developing strong relationships may prove difficult for firms that want to move customers from dedicated salespeople to offshore call centers for various reasons. These firms should recognize that a lack of seller expertise, dissimilarities between boundary spanners and customers, ineffective communication, and shifts from interpersonal to person–firm relationships can negatively affect customer–seller relationships.

Limitations

Meta-analyses have several strengths, but they also contain inherent limitations. First, the constructs we include are constrained to variables for which sufficient primary data are available. Thus, our framework should be considered a summary of the most commonly studied RM constructs, not an exhaustive list or even a list of the most important constructs. For example, mutual dependence, seller disclosure, and functional conflict have been shown to be important constructs for RM, but because of data unavailability, we could not include them in our meta-analysis. Second, heterogeneity in effect sizes remained even after we accounted for any variability due to the moderator variables in the study, which indicates that the effect sizes we report should be considered averages and may vary with the inclusion of unmeasured moderating conditions. Third, because of the limited number of studies for some moderator variables, our study has limited power to reject null hypotheses.

Future Research Directions

After nearly two decades of RM research, marketers' efforts may need to shift from significant testing to identifying which, and in what conditions, RM strategies generate the highest return on RM investment. Our synthesis of the extant literature identifies several avenues that require further study.

Research should expand the constructs included in our RM-mediated framework and determine which aspects or dimensions should be included to obtain a multifaceted view of relational exchanges. Although commitment and trust play critical roles, other candidates might include relationship satisfaction, exchange efficiency, equity, relational norms, and reciprocity. The absence of any measure of reciprocity between exchange partners is especially notable because it has been identified as "the core of marketing relationships" (Bagozzi 1995, p. 275) and may help explain the pattern of effects surrounding the impact of relationship investments and benefits on relational mediators. Integrating reciprocity into the relational-mediating framework may also explain the large, direct effect of relationship investment on performance, such that people's inherent desire to repay "debts" generated by sellers' investments may lead to performance-enhancing behaviors, independent of trust or commitment.

In addition to taking a multidimensional perspective of relationships, the scope of RM research should expand to investigate potential interactions among the relational mediators and identify relational synergies. For example, the strong linkage between relationship quality and objective performance may be due to interactions among the different facets of a relationship.

Even some of the high-impact antecedents and important outcomes in our framework appear in relatively few primary studies (i.e., conflict, seller expertise, and WOM), which suggests the need for additional research. Seller expertise, beyond product-specific expertise, might include overall customer knowledge, industry expertise, creativity, process knowledge, and intraorganizational facilitation. Strategies that remedy conflict-laden events, such as service or relationship recoveries, also are critical to incorporate into both practice and further research.

The relatively small correlations between customer focal antecedents (relationship benefit, dependence on seller) and relational mediators are surprising because we took the relational mediators from the customer's perspective as well. This finding may be due to a misspecification; we may not have studied some critical customer-focused antecedents. Thus, researchers should investigate other customer-focused antecedents, such as perceived exchange efficiency, perceived relationship investments, and liking, to identify other key drivers of a strong relationship from the customer's perspective.

The heterogeneity across nearly all linkages, even after we account for the moderators we included, demands research to determine other moderators that may influence RM effectiveness (e.g., relationship age, customer control, customer involvement, relationship orientation of the customer). For example, as the customer's need for a relationship increases, RM strategies may become more effective. Thus, researchers should develop a measure of the relationship orientation of the customer to support the segmentation of RM efforts. Marketers could then target their RM efforts toward customers with the highest susceptibility for RM. Contrary to most existing RM research, our results and social psychology theory suggest that researchers should differentiate between individual–individual and individual–firm relationships.

In summary, we provide insight into the most effective RM strategies, the conditions that moderate this effectiveness, and how the links between both antecedents and consequences of relational mediators depend on the mediator being investigated. These insights provide managers with opportunities to improve the returns on their RM investments and researchers with directions to build more robust models of the influence of RM on outcomes.

REFERENCES

Ajzen, Icek and Martin Fishbein (1980), *Understanding Attitudes and Predicting Social Behavior.* Englewood Cliffs, NJ: Prentice Hall.

Anderson, Erin and Barton A. Weitz (1989), "Determinants of Continuity in Conventional Industrial Channel Dyads," *Marketing Science*, 8 (Fall), 310–23.

——— and ——— (1992), "The Use of Pledges to Build and Sustain Commitment in Distribution Channels," *Journal of Marketing Research*, 29 (February), 18–34.

Anderson, James C. and James A. Narus (1990), "A Model of Distributor Firm and Manufacturer Firm Working Partnerships," *Journal of Marketing*, 54 (January), 42–58.

——— and ——— (1991), "Partnering as a Focused Market Strategy," *California Management Review*, 33 (Spring), 95–113.

——— and ——— (2004), *Business Market Management: Understanding, Creating, and Delivering Value*. Upper Saddle River, NJ: Prentice Hall.

Bagozzi, Richard P. (1995), "Reflections on Relationship Marketing in Consumer Markets," *Journal of the Academy of Marketing Science*, 23 (4), 272–77.

Berry, Leonard L. (1996), "Retailers with a Future," *Marketing Management*, 5 (Spring), 39–46.

Brown, Steven P (1996), "A Meta-Analysis and Review of Organizational Research on Job Involvement," *Psychological Bulletin*, 120 (2), 234–49.

——— and Robert A. Peterson (1993), "Antecedents and Consequences of Salesperson Job Satisfaction: Meta-Analysis and Assessment of Causal Effects," *Journal of Marketing Research*, 30 (February), 63–77.

Cheung, Mike W. and Wai Chan (2005), "Meta-Analytic Structural Equation Modeling: A Two-Stage Approach," *Psychological Methods*, 10 (1), 40–64.

Cheung, Shu Fai and Darius K.S. Chan (2004), "Dependent Effect Sizes in Meta-Analysis: Incorporating the Degree of Interdependence," *Journal of Applied Psychology*, 89 (5), 780–91.

Cohen, Jacob (1977), *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press.

Colgate, Mark R. and Peter J. Danaher (2000), "Implementing a Customer Relationship Strategy: The Asymmetric Impact of Poor Versus Excellent Execution," *Journal of the Academy of Marketing Science*, 28 (3), 375–87.

Crosby, Lawrence A., Kenneth R. Evans, and Deborah Cowles (1990), "Relationship Quality in Services Selling: An Interpersonal Influence Perspective," *Journal of Marketing*, 54 (July), 68–81.

Devine, Patricia G. (1995), "Prejudice and Out-Group Perception," in *Advanced Social Psychology*, Abraham Tesser, ed. New York: McGraw-Hill, 467–524.

De Wulf, Kristof, Gaby Odekerken-Schröder, and Dawn Iacobucci (2001), "Investments in Consumer Relationships: A Cross-Country and Cross-Industry Exploration," *Journal of Marketing*, 65 (October), 33–50.

Doney, Patricia M. and Joseph P. Cannon (1997), "An Examination of the Nature of Trust in Buyer–Seller Relationships," *Journal of Marketing*, 61 (April), 35–51.

Dwyer, Robert F., Paul H. Schurr, and Sejo Oh (1987), "Developing Buyer–Seller Relationships," *Journal of Marketing*, 51 (April), 11–27.

Fern, Edward F. and Kent B. Monroe (1996), "Effect Size Estimates: Issues and Problems in Interpretation," *Journal of Consumer Research*, 25 (September), 89–105.

Furlow, Carolyn F. and S. Natasha Beretvas (2005), "Meta-Analytic Methods of Pooling Correlation Matrices for Structural Equation Modeling Under Different Patterns of Missing Data," *Psychological Methods*, 10 (2), 227–54.

Ganesan, Shankar (1994), "Determinants of Long-Term Orientation in Buyer–Seller Relationships," *Journal of Marketing*, 58 (April), 1–19.

Grewal, Dhruv, Sukumar Kavanoor, Edward F. Fern, Caroly Costley, and James Barnes (1997), "Comparative Versus Noncomparative Advertising: A Meta-Analysis," *Journal of Marketing*, 61 (October), 1–15.

Gruen, Thomas W., John O. Summers, and Frank Acito (2000), "Relationship Marketing Activities, Commitment, and Membership Behaviors in Professional Associations," *Journal of Marketing*, 64 (July), 34–49.

Gundlach, Gregory T., Ravi S. Achrol, and John T. Mentzer (1995), "The Structure of Commitment in Exchange," *Journal of Marketing*, 59 (January), 78–92.

Hamilton, David L. and Steven J. Sherman (1996), "Perceiving Persons and Groups," *Psychological Review*, 103 (2), 336–55.

Hennig-Thurau, Thorsten, Kevin P. Gwinner, and Dwayne D. Gremler (2002), "Understanding Relationship Marketing Outcomes: An Integration of Relational Benefits and Relationship Quality," *Journal of Service Research*, 4 (February), 230–47.

Hibbard, Jonathan D., Frederic F. Brunel, Rajiv P. Dant, and Dawn Iacobucci (2001), "Does Relationship Marketing Age Well?" *Business Strategy Review*, 12 (4), 29–35.

———, Nirmalya Kumar, and Louis W. Stern (2001), "Examining the Impact of Destructive Acts in Marketing Channel Relationships," *Journal of Marketing Research*, 38 (February), 45–61.

Hunter, John E. and Frank L. Schmidt (1990), *Methods of Meta-Analysis*. Newbury Park, CA: Sage Publications.

Iacobucci, Dawn and Amy Ostrom (1996), "Commercial and Interpersonal Relationships: Using the Structure of Interpersonal Relationships to Understand Individual-to-Individual, Individual-to-Firm, and Firm-to-Firm Relationships in Commerce," *International Journal of Research in Marketing*, 13 (1), 53–72.

Jap, Sandy D. and Shankar Ganesan (2000), "Control Mechanisms and Relationship Life Cycle: Implications for Safeguarding Specific Investments and Developing Commitment," *Journal of Marketing Research*, 37 (May), 227–45.

Johnson, Jean L. (1999), "Strategic Integration in Industrial Distribution Channels: Managing the Interfirm Relationship as a Strategic Asset," *Journal of the Academy of Marketing Science*, 27 (1), 4–18.

Kumar, Nirmalya, Lisa K. Scheer, and Jan-Benedict E.M. Steenkamp (1995), "The Effects of Supplier Fairness on Vulnerable Resellers," *Journal of Marketing Research*, 32 (February), 54–65.

Lagace, Rosemary R., Robert Dahlstrom, and Jule B. Gassenheimer (1991), "The Relevance of Ethical Salesperson Behavior on Relationship Quality: The Pharmaceutical Industry," *Journal of Personal Selling & Sales Management*, 11 (Fall), 39–47.

Mohr, Jakki J., Robert J. Fisher, and John R. Nevin (1996), "Collaborative Communication in Interfirm Relationships: Moderating Effects of Integration and Control," *Journal of Marketing*, 60 (July), 103–115.

Moorman, Christine, Gerald Zaltman, and Rohit Deshpandé (1992), "Relationships Between Providers and Users of Market Research: The Dynamics of Trust Within and Between Organizations," *Journal of Marketing Research*, 29 (August), 314–29.

Morgan, Robert M. and Shelby D. Hunt (1994), "The Commitment–Trust Theory of Relationship Marketing," *Journal of Marketing*, 58 (July), 20–38.

Oliver, Richard L. (1999), "Whence Consumer Loyalty?" *Journal of Marketing*, 63 (Special Issue), 33–44.

Orwin, Robert G. (1983), "A Fail-Safe *N* for Effect Size in Meta-Analysis," *Journal of Educational Statistics*, 8 (2), 157–59.

Reichheld, Fredrick F. (2003), "The One Number You Need," *Harvard Business Review*, 81 (December), 46–54.

Reinartz, Werner J. and V. Kumar (2003), "The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration," *Journal of Marketing*, 67 (January), 77–99.

Reynolds, Kristy E. and Sharon E. Beatty (1999), "Customer Benefits and Company Consequences of Customer-Salesperson Relationships in Retailing," *Journal of Retailing*, 75 (1), 11–32.

Rosenthal, Robert (1979), "The 'File-Drawer Problem' and Tolerance for Null Results," *Psychological Bulletin*, 86 (3), 638–41.

——— (1994), "Parametric Measures of Effect Size," in *The Handbook of Research Synthesis*, Harris Cooper and Larry V. Hedges, eds. New York: Russell Sage Foundation, 231–44.

Shadish, William R. and Keith C. Haddock (1994), "Combining Estimates of Effect Size," in *The Handbook of Research Synthesis*, Harris Cooper and Larry V. Hedges, eds. New York: Russell Sage Foundation, 261–84.

Shiv, Baba, Julie A. Edell, and John W. Payne (1997), "Factors Affecting the Impact of Negatively and Positively Framed Ad Messages," *Journal of Consumer Research*, 24 (December), 285–94.

Siguaw, Judy A., Penny M. Simpson, and Thomas L. Baker (1998), "Effects of Supplier Market Orientation on Distributor Market Orientation and the Channel Relationship: The Distributor Perspective," *Journal of Marketing*, 62 (July), 99–111.

Sirdeshmukh, Deepak, Jagdip Singh, and Barry Sabol (2002), "Consumer Trust, Value, and Loyalty in Relational Exchanges," *Journal of Marketing*, 66 (January), 15–37.

Smith, J. Brock (1998), "Buyer-Seller Relationships: Bonds, Relationship Management, and Sex-Type," *Canadian Journal of Administrative Sciences*, 15 (1), 76–92.

Spekman, Robert E. (1988), "Strategic Supplier Selection: Understanding Long-Term Relationships," *Business Horizons*, 31 (July–August), 75–81.

Srinivasan, Raji and Christine Moorman (2005), "Strategic Firm Commitments and Rewards for Customer Relationship Management in Online Retailing," *Journal of Marketing*, 69 (October), 193–200.

Sterne, Jonathan A.C. and Matthias Egger (2001), "Funnel Plots for Detecting Bias in Meta-Analysis: Guidelines on Choice of Axis," *Journal of Clinical Epidemiology*, 54, 1046–1055.

Szymanski, David M. and David H. Henard (2001), "Customer Satisfaction: A Meta-Analysis of the Empirical Evidence," *Journal of the Academy of Marketing Science*, 29 (1), 16–35.

Vargo, Stephen L. and Robert F. Lusch (2004), "Evolving to a New Dominant Logic for Marketing," *Journal of Marketing*, 68 (January), 1–17.

Zeithaml, Valarie A., A Parasuraman, and Leonard L. Berry (1985), "Problems and Strategies in Services Marketing," *Journal of Marketing*, 49 (Spring), 33–46.

Carolyn J. Simmons & Karen L. Becker-Olsen

Achieving Marketing Objectives Through Social Sponsorships

The corporate social responsibility literature has focused on creating broad associations, such as simply being "socially responsible," and on the simple transfer of positive affect from sponsored causes to sponsoring firms. Both views fail to recognize the power of social initiatives as a means for differentiating among socially responsible firms and, in particular, for reinforcing the brand's positioning. The authors adopt a more traditional branding perspective and show that the fit between a firm's specific associations and a sponsored cause can reinforce or blur the firm's positioning, influence liking for the sponsorship, and bolster or undermine the firm's equity. They also show that communications decisions can mitigate the negative effects of low fit. Finally, they show that sponsorship effects can persist for as long as a year despite day-to-day exposure to other brand communications.

Corporate support of social causes has emerged as a popular promotional tool. The logic is that consumers will view a firm more favorably if it supports social causes, and surveys suggest that this is true.[1] For example, 80% of respondents to a national survey report that social sponsorships create a more favorable impression of the firm; two-thirds of respondents also report that they would be more likely to purchase from a firm that supports social causes (Reilly 2000). However, a variety of theory and evidence indicate that the impact of social sponsorships may be more complex (Hoeffler and Keller 2002; Jagre, Watson, and Watson 2001; Lichtenstein, Drumwright, and Braig 2004; Pracejus and Olsen 2004; Rifon et al. 2004; cf. Sen and Bhattacharya 2001) and that when the sponsored cause is viewed as not fitting with the firm's image, sponsorships may even be harmful (Speed and Thompson 2000). We test this proposition and show that high-fit sponsorships (i.e., those for which partners are perceived as congruent on some key dimension) can increase brand equity, whereas low-fit sponsorships (i.e., those for which partners are perceived as incongruent) can dilute brand equity. Furthermore, we show that these effects occur because low fit blurs the firm's positioning and decreases the favorability of attitudes toward the sponsorship.

Recognizing that not all firms have a natural high fit with a chosen cause, we also show that the negative effects of low fit can be reduced (1) through communications decisions that create perceived fit and (2) by altering the message source. Creating fit—that is, making salient how the two organizations are congruent—may be as effective as a naturally high-fit cause, whereas communicating through the sponsored organization rather than through the firm can reduce the perception of low fit and its associated risks. Finally, we show that fit effects may persist for as long as a year.

Contribution

The idea that low fit between a brand and other entities with which it is associated (e.g., extensions, alliance partners) can be harmful is not a new one (e.g., Aaker 1990; Keller and Aaker 1992; John, Loken, and Joiner 1998; Milberg, Park, and McCarthy 1997). However, we make several contributions. First, it might be expected that intrinsically positive activities, such as social sponsorships, are less likely to suffer from low fit than are purely commercial activities, such as brand extensions or alliances. In contrast, we show that the benefits of intrinsically favorable actions can be mitigated and even reversed by low fit (cf. Hoeffler and Keller 2002). Second, the corporate social responsibility literature has focused on creating broad associations, such as being "socially responsible," and on the simple transfer of positive affect from sponsored causes (Brown and Dacin 1997; Rifon et al. 2004; Smith 2004; cf. Sen and Bhattacharya 2000). We adopt a more traditional branding perspective and show that the fit between a firm's specific associations and a sponsored cause can reinforce or blur the firm's positioning. We also show that the unexpectedness of low fit leads to increased elaboration on the sponsorship and that this elaboration is negatively biased, leading to less favorable attitudes toward the sponsorship. Third, we extend work on how message variables moderate the effects of fit (Bridges, Keller,. and Sood 2000), showing that message content and source can each reduce the unfavorable effects of low fit. Fourth, we offer the first experimental evi-

---

[1] Sponsors can be either firms or brands, but for simplicity, we use "firms."

Carolyn J. Simmons is a visiting associate professor, Williams School of Commerce, Economics, and Politics, Washington and Lee University (e-mail: simmonsc@wlu.edu). Karen L. Becker-Olsen is an assistant professor, School of Business, The College of New Jersey (e-mail: kbecker@tcnj.edu). This research was supported in part by a summer research grant from the University of Virginia. Much of the work was completed while Carolyn J. Simmons was a visiting scholar at Virginia Polytechnic Institute and State University. The authors thank David Mick, Rick Netemeyer, and the three anonymous *JM* reviewers for their comments on previous versions of this work.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

dence of the enduring effects of sponsorship, showing effects that endure for up to a year. Fifth, we add to the literature on sponsorship recall. Prior work has shown a bias in favor of recalling semantically related pairings (Johar and Pham 1999; Pham and Johar 2001). We show that communications that create fit between semantically unrelated organizations can lead to high levels of recall even after a one-year delay.

Theoretical Background

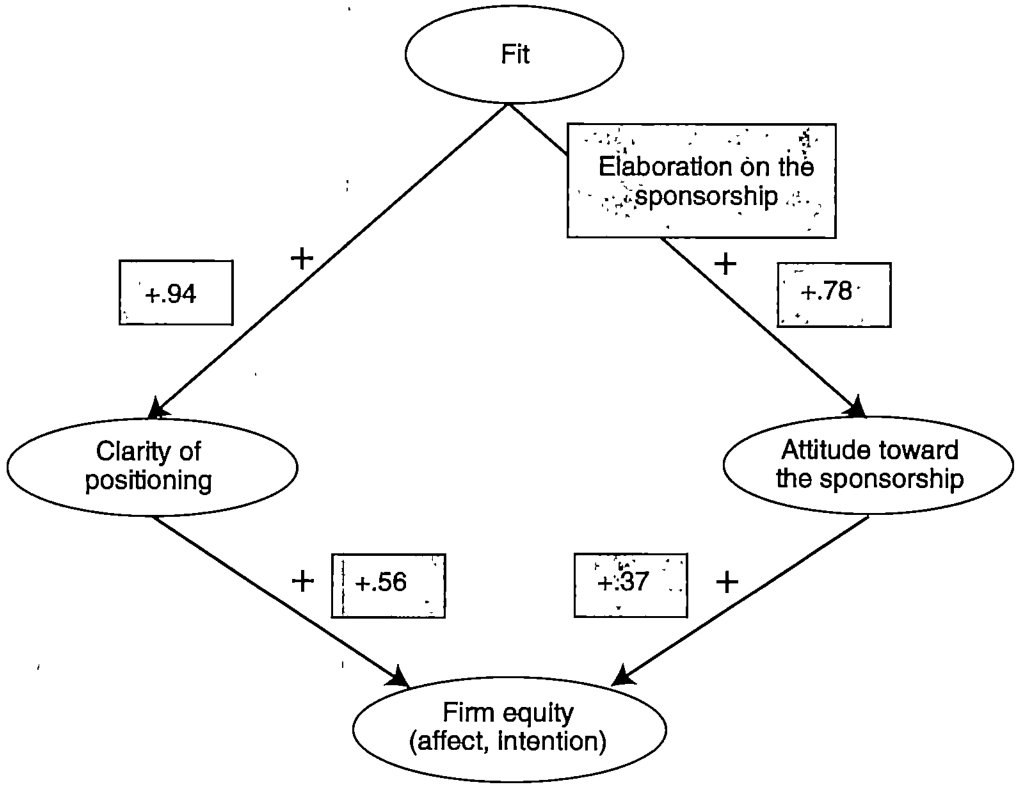
What Is Fit, and Why Does It Matter?

In the management literature, the fit construct embodies the idea of transferability of expertise or synergies in activities, such as when there is similarity in products, technologies, or markets (Rumelt 1974) or complementarity of skills and activities (Porter 1987). A similar conceptualization underlies Sen and Bhattacharya's (2001) examination of fit in the domain of internal social initiatives. In the brand extension and brand alliance literature, the focus is often on consumers' perceptions of fit. Although it is assumed that ideas of transferability and synergy often underlie perceived fit, it is recognized that what is transferred or enhanced may be intangible associations (Aaker and Keller 1990; Broniarczyk and Alba 1994; John, Loken, and Joiner 1998; Lichtenstein, Drumwright, and Braig 2004; Loken and John 1993; Simonin and Ruth 1998). We adopt a similar consumer-based conceptualization. Fit between a firm and a sponsored cause is high when the two are perceived as congruent (i.e., as going together), whether that congruity is derived from mission, products, markets, technologies, attributes, brand concepts, or any other key association (Bridges, Keller, and Sood 2000; Park, Milberg, and Lawson 1991).

Fundamentally, fit matters because high-fit sponsorships are consistent with what would be expected from the firm, whereas low-fit sponsorships are not. As we discuss in greater detail subsequently, this difference affects the clarity of the firm's positioning, and it alters processing of sponsorship information to influence how well the sponsorship is liked. In turn, clarity and attitude toward the sponsorship influence affective and behavioral responses to the firm, two key components of equity (Keller 1993). We outline these propositions in Figure 1.

Clarity of positioning. Consistency in communication helps build clarity, or the extent to which people know what to expect from the firm (Erdem and Swait 1998; Keller 1993; Park, Jaworksi, and MacInnis 1986). Because high-fit

FIGURE 1
Conceptual Framework and Study 1 Path Estimates (In Boxes)


Fit
Elaboration on the sponsorship
+
+.94
+
+.78
Clarity of positioning
Attitude toward the sponsorship
+
+.56
+.37
+
Firm equity (affect, intention)

Notes: Path estimates are standardized regression coefficients (b*). All *p*s < .001. We also examined elaboration as a mediator of the effect of fit on attitude toward the sponsorship. We find that the effect of fit on attitude is partially mediated by elaboration on the sponsorship ($b^*_{\text{fit, elaboration}} = .49$, $p < .001$; $b^*_{\text{elaboration, attitude}} = .15$, $p < .01$, and $b^*_{\text{fit, attitude}} = .70$, $p < .001$). The direct effect of fit may occur because low fit is negatively valued; however, this result may simply reflect that our thought measures do not capture intensity in the way that our scaled measures do (i.e., thoughts are coded as positive, neutral, or negative).

sponsorships are consistent with what people expect from the firm, such sponsorships should increase clarity. Because low-fit sponsorships are inconsistent with what people expect from the firm, such sponsorships should reduce clarity.

Attitude toward the sponsorship. On its surface, a social sponsorship is an intrinsically favorable action, one that we expect consumers to like. However, the interpretation of marketing activities, as is the case with the interpretation of any behavior, often involves two stages (Campbell and Kirmani 2000; Gilbert 1989; Johar and Simmons 2000). People first make a relatively effortless inference based on a surface interpretation (e.g., a salesperson's flattering remark reflects his or her true beliefs, a product with an unfamiliar brand name is of low quality). Then, and only if they allocate sufficient processing capacity, people may "correct" this inference through a more effortful process that accounts for other accessible inputs (e.g., the salesperson has an interest in ingratiating him- or herself with the customer, an unfamiliar brand is made by a well-respected firm). Similarly, we expect that a social sponsorship will be interpreted according to its surface meaning—that is, as a good act that is worthy of liking—unless people engage in further elaboration, in which case they may consider any other accessible inputs to judgment. Low fit plays a dual role in this process: (1) It increases cognitive elaboration, and (2) it makes countering (negative) inputs accessible.

Why would these effects occur? Low fit should increase elaboration because unexpected events engender thought, often in the form of an attributional search to understand the reasons for the event (Weiner 1985). Low fit should make countering inputs accessible because low fit (i.e., incongruity, unexpectedness) is negatively valued (Mandler 1982). Through a process of attitude priming (Fazio and Williams 1986; Houston and Fazio 1989), this negative affect should engender other negative thoughts, leading to a less favorable attitude toward the sponsorship. Similar effects of fit on attitudes have been found for brand alliances and extensions (Aaker and Keller 1990; Boush and Loken 1991; Broniarczyk and Alba 1994; Simonin and Ruth 1998), but in neither of these domains have increased elaboration and specific negative cognitions been demonstrated.

Firm equity. In turn, clarity of positioning and attitude toward the sponsorship should influence affective and behavioral responses to the firm, two key components of equity (Keller 1993). Clarity should have an effect because it alters the value of the brand as a signal. When clarity is high, buyers who value the brand's benefits perceive less risk in choosing the brand and feel less need to gather information to make that choice (Erdem and Swait 1998). Attitude toward the sponsorship should have an effect because it is a new affective association for the firm. Consistent with this prediction, responses to brands have been shown to be correlated with attitudes toward brand alliances (Simonin and Ruth 1998) and extensions (Erdem 1998; Sullivan 1990). In summary, we expect that the firm's equity increases with high-fit sponsorships and decreases with low-fit sponsorships and that clarity of positioning and attitude toward the sponsorship mediate these effects.

Hypotheses

On the basis of the foregoing discussion, we propose the following two hypotheses:

> H_1: Low-fit (versus high-fit) sponsorships lead to (a) decreased clarity of positioning, (b) greater cognitive elaboration (i.e., more thoughts), (c) less favorable thoughts, (d) a less favorable attitude toward the sponsorship, and (e) reduced equity, as measured by affective and behavioral responses to the firm.
>
> H_2: Clarity of positioning and attitude toward the sponsorship mediate the effect of fit on equity.

Study 1: Natural Fit

We tested these hypotheses in an experiment in which we asked respondents to read a press clipping that contained four business news items. One item reported on the opening of an expanded online store by one of two firms and either did or did not mention the firm's sponsorship of a nonprofit organization. When a sponsorship was mentioned, the sponsored organization was either high or low in fit with the company's image, as revealed by pretests. After respondents read these items, they responded to our dependent measures.

Because we are interested in the ability of sponsorships both to dilute and to enhance equity, we focus on familiar, well-liked firms. Furthermore, we focus on the effects of "natural fit"—that is, the extent to which the sponsored cause is perceived as being congruent with the image of the sponsor, independent of efforts to create a perceived fit between the organizations. For example, Alpo and the Humane Society are high in natural fit because both are strongly associated with pets; this association readily comes to mind. In contrast, Alpo and the Special Olympics are low in natural fit because they share no highly accessible association. Alpo is associated with pet food, and the Special Olympics is associated with athletic competitions for the mentally disabled.

Natural fit is of interest for several reasons. First, firms may engage in low-fit sponsorships because of a sincere interest in the cause or a belief that fit is irrelevant. Second, cost efficiencies are greater if there is no need to spend on efforts to create fit. Finally, because marketers do not control all the contexts in which consumers encounter information about their activities, a sponsorship that does not depend on such control for its effectiveness is highly attractive.

Method

Pretests. We attempted to identify a pair of equally well-liked, familiar companies (C_1 and C_2) that we could pair with a set of equally well-liked, familiar nonprofits (N_1 and N_2), such that for C_1, N_1 would be a high fit and N_2 would be a low fit, whereas for C_2, N_1 would be a low fit and N_2 would be a high fit. Thus, our fit manipulation would not be confounded with any differences between nonprofits.

Alpo and Sports Authority paired with the Humane Society and the Special Olympics met our criteria. The two firms were equally well liked (Ms = 6.10 and 6.17; $F_{1,27} < 1$) and familiar (Ms = 6.50 and 6.57; $F_{1,27} < 1$); we measured both liking and familiarity with three (seven-point) semantic differential items: "negative/positive," "unfavorable/favorable," and "bad/good" (Cronbach's α = .97), and "unfamiliar/familiar," "did not recognize/recognized," and "had not heard of/had heard of" (Cronbach's α = .96). Furthermore, the two nonprofits were also equally well liked (Ms = 6.25 and 6.36; $F_{1,21} = 1.63$, $p > .21$) and familiar (Ms = 6.13 and 6.15; $F_{1,21} < 1$). Fit, which we measured as "dissimilar/similar," "inconsistent/consistent," "atypical/typical," "unrepresentative/representative," "not complementary/complementary," "low fit/high fit," and "does not make sense/makes sense" (Cronbach's α = .99), varied as we expected (Ms = 6.45 and 1.80; $F_{1,25} = 500.09$, $p < .0001$), and this effect did not differ by firm ($F < 1$). Furthermore, fit did not differ between firms for high-fit nonprofits (Ms = 6.65 and 6.25; $F_{1,25} = 1.57$, $p > .22$) or for low-fit nonprofits (Ms = 1.84 and 1.75; $F_{1,25} < 1$).

Respondents and design. Two hundred thirty-six students in nonmarketing professional seminars participated in the study without compensation. Their average age was 35, and 52% were male. The design is a 2 (fit: high versus low) × 2 (sponsorship: present versus absent) × 2 (company: Alpo versus Sports Authority) × 2 (measurement order: sponsorship questions before versus after equity measures) completely randomized between-subjects design. This design provides different controls for the low-fit and high-fit conditions (sponsorship: present versus absent is crossed with fit) to ensure that we isolate the effects of sponsorship per se. Particularly in the high-fit condition, simply thinking about the nonprofit and the firm at the same time may lead to more favorable thoughts about the firm, for example, being exposed to the Humane Society might make a person think about how much he or she likes dogs, which may make this person think more favorably about a firm such as Alpo, which produces dog food. Our control conditions isolate this effect from the effect of sponsorship because they expose respondents to both organizations in the absence of a sponsorship relationship. The measurement order factor allows us to determine whether effects on firm equity occur spontaneously or whether they occur only if consumers are directed to think about and evaluate the sponsorship.

Materials. The materials we used appear in Appendix A. Each clipping contains four news items. The first two items are constant across conditions. The third item varies across conditions and describes the promotion of an employee of either the Humane Society or the Special Olympics. The last item also varies across conditions and describes a new Web site for either Alpo or Sports Authority. In the sponsorship condition, it also announces the sponsorship program. There is a no-sponsorship control for every pairing of a company with a nonprofit. The only difference between no-sponsorship and sponsorship conditions is the presence of the sponsorship.

Measures. We measured fit as we did in the pretest. We measured elaboration by asking respondents to list all the thoughts they had while thinking about the sponsorship. Two coders, who were blind to the conditions, categorized thoughts as thoughts about fit (e.g., "The organizations seem to go together," "Pet food and the Special Olympics have nothing to do with each other"), the sponsorship (e.g., "This is a great thing," "It is actually a little offensive"), efficacy for the firm ("They may open more of a market for themselves," "I feel that Alpo will not sell more food"), the firm's motivation (e.g., "Seems as though Alpo is trying to help," "What a cry for publicity"), and other thoughts (e.g., "My dog eats Alpo," "I like pets"). Fit thoughts were further classified as high, uncertain, or low fit. Inferences about the firm's motivation were classified as social, uncertain, or self-interested motivation. All other thoughts were classified as favorable, neutral, or unfavorable. Agreement between coders was 81%, and there were no significant differences in their coding for any thought category ($ps > .45$). Discrepancies were resolved by discussion.

We measured clarity of positioning by rated agreement (seven-point scales) with the items "clearly communicates what it stands for," "has an image that is difficult to understand," and "conveys a clear image in all its actions" (Cronbach's α = .94). We measured attitude toward the sponsorship with the same attitudinal items used in the pretest for companies and nonprofits. We measured firm equity with these same three attitudinal measures and three behavioral intention measures, the rated likelihood to "consider," "purchase," and "recommend" (Cronbach's α = .96).[2]

Procedure. We told respondents that we were interested in their responses to news stories. They read the clipping and then responded to our dependent measures. We presented questions about the sponsorship and the firm in counterbalanced order, followed by all other measures. We then debriefed respondents and told them that the articles were fictitious.

Results

The results conform to our predictions. When fit is low versus high, clarity of positioning is lower, more thoughts are generated in response to the sponsorship, these thoughts are less favorable, attitude toward the sponsorship is less favorable, and firm equity (operationalized as affective and behavioral responses to the firm) is lower. Furthermore, clarity and attitude toward the sponsorship mediate the effect of fit on firm equity. Except where we subsequently note, none of these effects vary as a function of measurement order ($ps > .11$) or company ($ps > .30$). We present means in Table 1.

Manipulation check. Perceived fit varies as we expected (Ms = 1.59 and 6.45; $F_{1,152} = 3574.97$, $p < .0001$), and it does not vary by firm ($F_{1,152} = 1.83$, $p > .17$). Furthermore, perceived fit does not vary by exemplar in the high-fit condition (Ms = 6.48 and 6.42; $F_{1,76} < 1$) or in the low-fit condition (Ms = 1.51 and 1.67; $F_{1,76} = 1.61$, $p > .20$).

---

[2] Because not everyone buys pet supplies or sports-related products, we asked these questions as follows: "If you or a friend had the need to buy the kind of products sold by (Alpo, Sports Authority), how likely would you be to...?"

TABLE 1
Study 1 Means

	Alpo				Sports Authority				Combined			
	Low Fit		High Fit		Low Fit		High Fit		Low Fit		High Fit	
	Control	Sponsorship	Control	Sponsorship	Control	Sponsorship	Control	Sponsorship	Control	Sponsorship	Control	Sponsorship
Scaled Measures												
Clarity of positioning	5.20	2.78	5.15	6.15	5.10	2.94	5.03	6.20	5.15^1	2.86^{a2}	5.09^1	6.17^{b2}
Attitude toward the sponsorship	—	4.10	—	6.43	—	4.23	—	6.61	—	4.17^a	—	6.52^b
Firm equity	5.62	4.88	5.48	6.52	5.59	4.97	5.77	6.47	5.60^1	4.93^{a2}	5.62^1	6.50^{b2}
Number of Thoughts About ...												
Fit		1.15		.23		.64		.28		.89^a		.25^b
Sponsorship		1.08		.76		1.12		.81		1.10^a		.79^b
Inferred efficacy		.41		.30		.23		.15		.32		.23
Inferred motivation		.22		.05		.20		.21		.21		.13
Other		.84		.53		1.15		.78		1.00		.65
Total		3.72		1.89		3.35		2.24		3.53^a		2.07^b
Favorability of Thoughts About ...												
Fit		–.64		.23		–.51		.15		–.57^a		.19^b
Sponsorship		.25		.38		–.02^a		.58^b		.11^a		.48^b
Inferred efficacy		–.18		.30		–.10		.13		–.14^a		.21^b
Inferred motivation		–.12^a		.05^b		.10^a		–.10^b		–.01		–.02
Other		.49		.22		.41		.28		.45		.25
Total		–.20		1.20		–.13		1.04		–.16^a		1.12^b

Notes: High-fit and low-fit means with different alphabetical superscripts differ from each other ($p < .0001$). Control and sponsorship means with different numerical superscripts differ from each other ($p < .001$). When effects generalize across companies (no interaction with company), we summarize statistical tests with the use of alphabetical superscripts in the "Combined" column. When effects do not generalize across companies, we also summarize simple effects tests with the use of alphabetical superscripts in each company column.

Clarity of positioning. As H_{1a} predicted, when fit is low rather than high, respondents perceive the firm's positioning as less clear (Ms = 2.86 and 6.17; $F_{1,\,151} = 744.70$, $p < .0001$). Furthermore, control group comparisons show that clarity is lower with a low-fit sponsorship than with no sponsorship (Ms = 2.86 and 5.15; $F_{1,\,113} = 205.75$, $p < .0001$) and is higher with a high-fit sponsorship than with no sponsorship (Ms = 6.17 and 5.09; $F_{1,\,114} = 91.98$, $p < .0001$).

Number and favorability of thoughts. As H_{1b} predicted, more thoughts are generated when fit is low than when it is high (Ms = 3.53 and 2.07; $F_{1,\,147} = 34.88$, $p < .0001$). Because low fit is unexpected, we expect that low fit is noted more often than high fit, but the overall increase in thoughts is not attributable only to thoughts about low fit. When we omit fit thoughts, the total number of thoughts remains higher with low fit (Ms = 2.64 and 1.82; $F_{1,\,147} = 12.84$, $p < .0005$). An examination of individual thought categories shows that though low fit leads to more thoughts about fit (Ms = .89 and .25; $F_{1,\,147} = 32.76$, $p < .0001$), it also leads to more thoughts about the sponsorship (Ms = 1.10 and .79; $F_{1,\,147} = 4.60$, $p < .03$). Thus, the increased elaboration associated with low fit is focused not solely on fit but on other aspects of the sponsorship as well. No other effects on the number of thoughts are significant.

As H_{1c} predicted, thoughts are less favorable overall when fit is low than when it is high (Ms = –.16 and 1.12; $F_{1,\,147} = 39.65$, $p < .0001$).[3] Furthermore, this difference is not attributable solely to thoughts about fit. Thoughts remain less favorable with low fit when we omit fit thoughts (Ms = .41 and .93; $F_{1,\,147} = 7.54$, $p < .006$). Consistent with the idea of attitude priming, individual thought categories reveal relatively little else systematic other than the overall pattern of less favorable thoughts with low fit. Low fit leads to less favorable thoughts about fit (i.e., more thoughts about low fit; Ms = –.57 and .19; $F_{1,\,147} = 59.08$, $p < .0001$), the sponsorship (Ms = .11 and .48; $F_{1,\,147} = 9.11$, $p < .003$), and the efficacy of the sponsorship for the firm (Ms = –.14 and .21; $F_{1,\,147} = 20.93$, $p < .0001$). This last difference is marginally greater for Alpo ($F_{1,\,147} = 2.69$, $p < .10$) but is significant for both companies (ps < .02). The effect on sponsorship thoughts varies by company ($F_{1,\,147} = 3.77$, $p < .05$) and holds only for Sports Authority ($p < .001$) but not for Alpo ($p > .43$). Thoughts about motivation vary by company ($F_{1,\,147} = 10.08$, $p < .001$), becoming less favorable (i.e., more about self-interest) with low fit for Alpo ($F_{1,\,74} = 5.86$, $p < .01$) but more favorable (i.e., more about social motivation) with low fit for Sports Authority ($F_{1,\,73} = 4.59$, $p < .03$).

Attitude toward the sponsorship. As H_{1d} predicted, when fit is low rather than high, attitude toward the sponsorship is less favorable (Ms = 4.17 and 6.52; $F_{1,\,148} = 195.06$, $p < .0001$). The effect is marginally greater when the attitude measure is taken before the firm equity measure ($F_{1,\,148} = 2.72$, $p < .10$), but it is significant in either case (ps < .0001).

Firm equity. As H_{1e} predicted, when fit is low rather than high, the firm's equity is lower (Ms = 4.93 and 6.50; $F_{1,\,148} = 260.21$, $p < .0001$). Furthermore, control-group comparisons show that firm equity is lower with a low-fit sponsorship than with no sponsorship (Ms = 4.93 and 5.60; $F_{1,\,114} = 15.12$, $p < .0002$) and is higher with a high-fit sponsorship than with no sponsorship (Ms = 6.50 and 5.62; $F_{1,\,114} = 72.99$, $p < .0001$). The effect of a high-fit sponsorship on firm equity is marginally greater for Alpo ($F_{1,\,114} = 2.80$, $p < .09$), but it is significant for both firms (ps < .0001).

Mediation. H_2 predicted that the effect of fit on equity is mediated by clarity of positioning and attitude toward the sponsorship. Our measurement model includes four latent variables: fit, clarity, attitude toward the sponsorship, and firm equity. Fit is represented by three composite measures, which we constructed by randomly dividing our seven items into one three-item parcel and two two-item parcels (Kishton and Widamon 1994; Marsh et al. 1998). Clarity and attitude toward the sponsorship are represented by our three single-item measures in each case. Firm equity is represented by two composite measures, our three-item measure of attitude toward the firm and our three-item measure of behavioral intention toward the firm.

The measurement model provides a strong fit to our data. Although the chi-square is significant (65.79, d.f. = 38, $p < .003$), the comparative fit index (.99; Bentler 1989) and the nonnormed fit index (.98; Bentler and Bonett 1980) are both acceptably high. Composite reliabilities for the latent factors all exceed .91, indicating high internal consistency (Hatcher 1994, adapted from the work of Fornell and Larcker 1981). Variance extracted estimates for the latent factors all exceed .84, indicating a high amount of variance captured by the factors, compared with error variance (Fornell and Larcker 1981). Furthermore, all factor loadings are high (>.87) and highly significant ($p < .001$), indicating convergent validity. Although the factors are highly correlated, as our model predicts, discriminant validity is also good.[4] All factors exhibit discriminant validity, as a chi-square difference test (ps < .001) and a confidence interval test (ps < .001; Anderson and Gerbing 1988) indicate. A variance extracted test shows discriminant validity for all pairs of variables except fit and clarity, for which the squared correlation between factors (.89) slightly exceeds the variance extracted for clarity (.86). Thus, only 1 of 18 tests fails to show good discrimination; this is about the rate that would be expected to occur by chance at the standard alpha of .05. Furthermore, the one exception is for the method (variance extracted) that does not employ a significance test. For the other two methods, we employ the highly stringent alpha of .001.

---

[3] We coded thoughts about low fit and self-interested motivation as unfavorable thoughts.

[4] Factor correlations:

	Clarity	*Attitude Toward the Sponsorship*	*Equity*
Fit	.94	.78	.83
Clarity	—	.74	.84
Attitude toward the sponsorship	—	—	.79

Our theoretical model, which we illustrate with the ovals and paths in Figure 1, also exhibits strong fit. Although the model chi-square is significant (66.75, d.f. = 40, $p < .005$), both the comparative fit index (.99) and the nonnormed fit index (.98) indicate acceptably high fit. Each standardized path coefficient (b*) is nontrivial (>.36) and significant (*p*s < .001). Furthermore, the R-square for all latent endogenous variables is sizable (.89, .61, and .77 for clarity, attitude toward the sponsorship, and firm equity, respectively). Finally, clarity and attitude toward the sponsorship completely mediate the effect of fit. When a direct path from fit is added to the model, this path is not significant (b* = .16, $p > .332$), and model fit is not improved ($\chi^2 = .93, p > .334$).[5]

When we measure all variables with scaled measures collected at the same point in time, common methods variance is a potential problem (Doty and Glick 1998). Although we cannot completely rule out common methods variance, we use methods to limit such bias. First, we vary measurement order, so that respondents do not always evaluate firm-related measures before sponsorship-related measures, or vice versa. Second, we use thought data to validate our process model, as described previously. Finally, in a regression-based mediation analysis (Kenny, Kashy, and Bolger 1998), we employ a nonscaled measure of fit (i.e., our fit manipulation), and again, we find that the effect of fit on equity is mediated by clarity of positioning ($z = 2.90, p < .003$) and attitude toward the sponsorship ($z = 4.58, p < .001$). Because this analysis employs our fit manipulation, it also confirms fit as a causal variable within the context of the mediation analysis.

Discussion

A social sponsorship can be an effective way to enhance the firm's equity; however, a poor choice of a cause can eliminate and even reverse this effect, creating actual harm, as is shown by our comparisons with a no-sponsorship control group. Low fit decreases clarity, leads to negative elaboration on the sponsorship and a less favorable attitude toward the sponsorship, and dilutes the firm's equity, as measured by affective and behavioral responses. These results have generality across exemplars and measurement contexts. The same firm benefits or is harmed, depending on which of two equally well-liked and familiar nonprofits (i.e., high fit or low fit) it sponsors. The same nonprofit organization has either a favorable or an unfavorable effect, depending on which of two equally well-liked and familiar firms (i.e., high fit or low fit) is its sponsor. The effects of fit on thoughts, attitude toward the sponsorship, and firm equity occur regardless of whether equity or sponsorship measures are taken first; thus, these effects do not depend on thoughts engendered by previous measures (Feldman and Lynch 1988). Furthermore, our thought data validate our process interpretation. Even without the prompting provided by scaled measures, people think about whether a sponsorship is consistent with the firm's image, they have more favorable thoughts when they perceive a sponsorship as high fit, and these effects vary with our fit manipulation.

Our mediation analysis confirms that fit affects firm equity through clarity and attitude toward the sponsorship. This result suggests that social sponsorships should work through both routes. A firm may gain a positive attitude toward a low-fit sponsorship by adding attractive elements (e.g., a celebrity spokesperson) but still blur the brand image and dilute equity if fit is low. Similarly, a firm may reinforce the brand image with a high-fit sponsorship but still dilute brand equity if the sponsorship is disliked for other reasons (e.g., heavy-handed promotional efforts).

Finally, our data offer insights into a long-standing idea about fit. In the domains of brand extensions and corporate social responsibility, it has been argued that low fit engenders attributions that the firm is motivated by a desire to gain market value rather than by a desire to deliver value to consumers (Aaker 1990; Keller and Aaker 1992; cf. Brown and Dacin 1997). Rifon and colleagues (2004) find support for this prediction in the context of a sponsorship of a health information Web site. In contrast, we do not obtain this effect. Our thought listings (i.e., undirected responses to the sponsorships) show no consistent effect of fit on attributed motivations. Notably, we also took a subsequent scaled measure of motivation and obtained the predicted effect, but this measure did not exhibit discriminant validity. Thus, our data suggest that scaled measures of motivation, or any specific belief related to the sponsorship, should be interpreted with caution. Although thought listings are imperfect as well, scaled measures prompt judgments about constructs (e.g., motivation) that might not have otherwise been salient for many respondents. When this occurs, respondents may construct a judgment (e.g., based on attitude priming) that they might not otherwise make (Feldman and Lynch 1988; Simmons, Bickart, and Lynch 1993). Thus, the use of alternative measures, such as undirected thoughts or true behavioral measures, to validate that proposed responses occur spontaneously can greatly improve our understanding of how people respond to social sponsorships.

Study 2: Created Fit, Message Source, and the Persistence of Effects

Many firms may not have a high natural fit with many causes or may have an ongoing relationship with a low-fit cause, particularly with the current interest in corporate social responsibility and the common practice of sponsoring a standard set of causes (e.g., education, housing, health care). We now examine strategies that may help counter the negative effects of low fit. Specifically, we consider how message content can be used to create perceived fit and how message source may reduce the risks associated with low fit. We also examine how well the effects of the sponsorship persist over a one-year period.

Created Fit

Bridges, Keller, and Sood (2000) describe a relational communication strategy for brand extensions that is intended to increase perceived fit. When salient parent-brand associations are not relevant to the extension category but nonsalient associations are, the relational strategy involves emphasizing the relevant nonsalient associations. In the context of social sponsorships, there are at least two ways that a nonsalient shared association can be made salient and thus create perceived fit: (1) with product-related donations that evoke a shared association and (2) with messages that explain how the firm fits with the cause. As an example of using product-related donations to create fit, the Ford Motor Company (2002) recently helped make the vintage Red Bus Fleet of Glacier National Park safe for continued use. Without this donation, there is no apparent link between Ford and national parks; with it, the idea that parks require transportation services becomes salient. As an example of using a message to create fit, DenTek Oral Care (2006), a sponsor of the American Diabetes Association, includes in its sponsorship communications the information that diabetes can lead to tooth decay, bad breath, dry mouth, and gum disease. Because many people may not know about diabetes-related dental problems, the sponsorship might otherwise seem to be a bad fit.

---

[5] We also tested a model in which the effect of fit on attitude toward the sponsorship is mediated by clarity. It yielded significantly poorer fit than our theoretical model: $\chi^2 = 67.73, p < .001$.

To maximize the probability of successfully creating perceived fit, we use both strategies to create fit for one of our low-fit sponsorships from Study 1, namely, Alpo's sponsorship of the Special Olympics. Specifically, the sponsorship announcement includes the information that the Special Olympics participants will receive pet-related gifts, and it also explains that caring for pets increases the self-esteem of mentally disabled children. We expect that if we are successful in creating fit, we will observe benefits similar to those of natural fit.

Message Source

People are less critical of messages from sources they regard as unbiased or disinterested (Artz and Tybout 1999; Birnbaum and Stegner 1979; Weiner, LaForge, and Goolsby 1990). We expect that people will perceive the firm as more self-interested than the sponsored nonprofit in communicating its sponsorship activity. Study 1's thought listings show that people are aware that sponsorships help achieve marketing goals; many respondents commented on the efficacy of the sponsorship for increasing sales, gaining publicity, and so forth. Although the sponsorship benefits the nonprofit as well, that organization apparently has less to gain from communicating that the firm has given it support. Therefore, we expect that sponsorships are evaluated less critically when the message source is the sponsored nonprofit than when it is the sponsoring firm.

Because the issue of source bias has received relatively little systematic investigation, we consider alternative hypotheses about the form of this effect. Birnbaum and Stegner (1979) examine source bias in a study in which respondents estimated the value of a used car on the basis of its blue book value and the source's estimate. When they perceived the source as self-interested, their estimates were lower than when they perceived the source as disinterested. That is, source bias had a main effect on judgments. Thus, we might expect that a nonprofit message source will have a main effect on the favorability of responses to the sponsorship (i.e., a positive effect regardless of fit).

In contrast, Artz and Tybout (1999; see also Wiener, LaForge, and Goolsby 1990) examine source bias as a moderator of responses to message content. People expect experts to speak in quantitative terms, whereas they expect nonexperts to speak in more qualitative terms. Messages that violate these expectations are less effective in creating positive attitudes but only if people perceive the source as self-interested. If people perceive the source as disinterested, they do not "penalize" the message when it violates their expectations. Thus, source bias moderates the effect of violated expectations on judgments. This finding is particularly relevant to our situation because low-fit sponsorships violate expectations, whereas high-fit sponsorships do not. Thus, we might expect a nonprofit source to increase the favorability of responses only when fit is low.

Persistence of Sponsorship Effects

Whether sponsorship effects endure is of crucial interest to sponsoring firms. We examine effects on firm equity and sponsorship recall after a one-year delay, during which respondents were exposed to brand messages only as they occurred routinely in their daily lives. Although our examination of the persistence of effects on firm equity is largely exploratory, we expect that sponsorship recall is greater with created fit than with low fit. Prior research has shown that recall of sponsorship partners is not only based on pure recollection but also incorporates constructive processes (Johar and Pham 1999). When attempting to recall a sponsorship, people use heuristics such as market prominence or semantic relatedness to help them identify sponsorship partners. A result of this partially constructive process is that people recall high-natural-fit sponsorships better than low-fit sponsorships. Because created fit involves creating a semantic link between otherwise unrelated organizations, we expect that people will recall created-fit sponsorships better than the same low-fit pairing. Both low fit and created fit begin with low-natural-fit pairings and thus engender elaboration, but the created-fit sponsorship gains an advantage from the creation of this associative link. Study 1's data suggest that this is not a common outcome of the elaboration for low-fit sponsorships. Only 6% of Study 1's low-fit respondents reported finding such a link.

Hypotheses

On the basis of the foregoing discussion, we propose the following three hypotheses:

H_3: Low-fit (versus created-fit) sponsorships lead to (a) decreased clarity of positioning, (b) a less favorable attitude toward the sponsorship, and (c) reduced equity, as measured by affective and behavioral responses to the firm.

H_4: A nonprofit message source (versus a company message source) leads to a more favorable response. This effect takes one of two forms: (a) greater clarity, a more favorable attitude toward the sponsorship, and higher firm equity with a nonprofit message source, regardless of fit, or (b) greater clarity, a more favorable attitude toward the sponsorship, and higher firm equity, but only when fit is low.

H_5: A created-fit sponsorship is better recalled than the same low-fit pairing

Method

Materials (see Appendix B) and procedures are similar to those of Study 1. The design is a 2 (fit: low versus created) × 2 (message source: company versus nonprofit) factorial with a separate-sample, no-sponsorship control. Because Study 1 showed high generality of effects across exemplars and no substantive effects of measurement order for the effects of interest, we examine effects for only one company and always take sponsorship measures before company measures. We selected Alpo as the company and the Special Olympics as the nonprofit. The low-fit condition is identical to the low-fit condition in Study 1. In the created-fit condition, the sponsorship includes a gift of a pet and pet food to the Special Olympics participants. The gift of a pet is intended to evoke a connection of Alpo to pets and children; its products nourish the pets that are part of many happy childhoods. The announcement also offers an explanation of how the organizations fit together, noting that caring for pets increases the self-esteem of the mentally disabled. We collected all measures, except sponsorship recall, just after exposure to the sponsorship. We measured firm equity again after a one-year delay; we collected recall at the same time. One hundred twenty students from nonmarketing professional training seminars participated in the study without compensation. Their mean age was 36, and 49% were male.

Results

In general, the results conform to our predictions. As we expected, created fit offers benefits similar to those of natural fit. Furthermore, a nonprofit message source (versus a company source) can increase clarity, the favorability of attitude toward the sponsorship, and firm equity. Evidence from our thought measures suggests that these effects are more reliable when fit is low. In addition, recall of the sponsored cause is greater with created fit than with low fit. We made no prediction about the persistence of effects, but we find that the effect of fit on firm equity persists over a one-year period.

Manipulation checks. As we expected, perceived fit is greater when fit is created than when it is low (Ms = 4.81 and 2.17; $F_{1, 76} = 125.98$, $p < .0001$); this effect is not moderated by message source ($p > .20$). Perceived fit is also greater with a nonprofit than with a company as the source (Ms = 4.09 and 2.89; $F_{1, 76} = 26.06$, $p < .0001$). Although we did not predict this effect, it is consistent with the idea that low-fit sponsorships are not judged as harshly with a nonprofit message source.

Clarity, attitude toward the sponsorship, and firm equity. We present means in Table 2. As H_3 predicted, clarity is lower, attitude toward the sponsorship is less favorable, and firm equity is lower with low fit than with created fit (Ms = 3.72 and 5.96; $F_{1, 76} = 119.69$, $p < .0001$; Ms = 3.87 and 6.15; $F_{1, 76} = 79.19$, $p < .0001$; and Ms = 4.77 and 6.25; $F_{1, 76} = 67.87$, $p < .0001$). As H_{4a} predicted, these judgments are more favorable with a nonprofit message source than with a company source (Ms = 5.43 and 4.25; $F_{1, 76} = 32.88$, $p < .0001$; Ms = 5.62 and 4.40; $F_{1, 76} = 22.96$, $p < .0001$; and Ms = 5.74 and 5.27; $F_{1, 76} = 6.79$, $p < .01$). For clarity, there is also a fit × source interaction ($F_{1, 76} = 6.15$, $p < .01$), but it is not of the exact form that H_{3b} predicted, namely, that the simple effect of source will be significant only for low fit. Instead, the effect is larger for low fit, but it is significant in any case ($F_{1, 38} = 27.32$, $p < .001$ and $F_{1, 38} = 6.92$, $p < .01$). No other fit × source interactions are significant.

We made no predictions with regard to control-group comparisons; however, clarity is greater with created fit than with no sponsorship (for company source: Ms = 5.63 and 5.19; $F_{1, 58} = 4.66$, $p < .03$; for nonprofit source: Ms = 6.30 and 5.19; $F_{1, 58} = 79.48$, $p < .0001$) and is lower with low fit than with no sponsorship (for company source: Ms = 2.88 and 5.19; $F_{1, 58} = 276.09$, $p < .0001$; for nonprofit source: Ms = 4.56 and 5.19; $F_{1, 58} = 6.80$, $p < .01$). Firm equity exhibits a similar pattern (created fit: for company source: Ms = 6.00 and 5.45; $F_{1, 58} = 5.13$, $p < .02$; for nonprofit source: Ms = 6.49 and 5.45; $F_{1, 58} = 27.68$, $p < .0001$; low fit: for company source: Ms = 4.55 and 5.45; $F_{1, 58} = 15.07$, $p < .0003$; for nonprofit source: Ms = 5.00 and 5.45; $F_{1, 58} = 3.97$, $p < .05$). Note that though a nonprofit source increases the favorability of these judgments, it never completely overcomes the negative effects of low fit.

Thought listings. Again, we use respondents' undirected thoughts to validate our process explanation. The effects of

TABLE 2
Study 2 Means

		Low Fit		Created Fit		Fit Means		Source Means	
	Control	Company Source	NPO Source	Company Source	NPO Source	Low	Created	Company	NPO
Scaled Measures									
Clarity of positioning	5.19^{1}	2.88^{a2}	4.56^{b2}	5.63^{a2}	6.30^{b2}	3.72^{a}	5.96^{b}	4.25^{a}	5.43^{b}
Attitude toward the sponsorship	—	3.15	4.60	5.65	6.65	3.87^{a}	6.15^{b}	4.40^{a}	5.62^{b}
Firm equity	5.45^{1}	4.55^{2}	5.00^{2}	6.00^{2}	6.49^{2}	4.77^{a}	6.25^{b}	5.27^{a}	5.74^{b}

Notes: We indicate significant main effects of fit and source by pairs of means with different alphabetical superscripts under the "Fit Means" and "Source Means" headings, respectively. When there is a fit × source interaction, we indicate any significant simple effects by pairs of means with different alphabetical superscripts under the headings "Low fit" and "Created fit." All simple-effect means that differ significantly from the no-sponsorship control mean have a different numerical superscript from the control mean. All $ps < .05$. NPO = nonprofit organization.

created fit substantively replicate the effects of natural fit from Study 1; however, we do not report these effects in greater detail. With regard to the effects of message source, thought listings are consistent with the idea that a nonprofit message source makes respondents less critical of low fit. Fit thoughts exhibit a fit × source interaction ($F_{1, 75} = 3.76$, $p < .05$); only when fit is low, a nonprofit source leads to marginally less unfavorable thoughts about fit than does a company source (Ms = –.10 and –.42; $F_{1, 37} = 2.99$, $p < .09$). For both number and favorability of thoughts about the sponsorship, there is a fit × source interaction ($F_{1, 75} = 4.66$, $p < .03$ and $F_{1, 75} = 3.73$, $p < .05$); only when fit is low, a nonprofit source leads to more thoughts and more favorable thoughts about the sponsorship (number: Ms = .42 and 1.20; $F_{1, 37} = 6.47$, $p < .01$; favorability: Ms = .05 and .60; $F_{1, 37} = 7.86$, $p < .008$); again, only when fit is low (fit × source: $F_{1, 75} = 4.34$, $p < .04$), there is a decrease in the favorability of *other* thoughts with a nonprofit source (Ms = .52 and .05; $F_{1, 37} = 8.44$, $p < .006$). Although this proposition is highly speculative, it may be that the negative attitude priming associated with low fit becomes diverted from the sponsorship when a nonprofit source engenders less critical processing. There are no other effects on number or favorability of thoughts.

Separate-sample tests. We conducted additional tests to increase our understanding of the created-fit and message source variables. Because the gift in the created-fit condition increases the economic value of the sponsorship, we conducted a separate-sample test to determine whether the results for created fit are attributable to the gift. We find that the gift has little or no effect on perceived fit, attitude toward the sponsorship, and firm equity, whereas the fit explanation (i.e., caring for pets increases self-esteem) accounts for most of the variance in these measures. For clarity, the gift has an effect only when the fit explanation is given. Thus, the effects of created fit are not attributable to increased value but rather to fit per se. The design is a 2 (fit explanation: present versus absent) × 2 (gift: present versus absent) factorial; we randomly assigned 76 respondents to conditions. Both the fit explanation and the gift increase perceived fit; however, the effect of the gift is much weaker (fit explanation: Ms = 2.84 and 4.39; $F_{1, 71} = 62.79$, $p < .0001$; gift: Ms = 3.42 and 3.81; $F_{1, 71} = 3.83$, $p < .05$). For attitude toward the sponsorship and firm equity, the gift has no impact (*ps* > .43), whereas the fit explanation increases favorability (attitude: Ms = 5.08 and 6.25; $F_{1, 71} = 16.30$, $p < .0001$; firm equity: Ms = 4.92 and 5.59; $F_{1, 71} = 10.01$, $p < .002$). For perceived clarity, there is a fit explanation × gift interaction ($F_{1, 70} = 4.10$, $p < .04$) such that the gift has an impact only if the fit explanation is given (no explanation: Ms = 3.90 and 4.08; $F_{1, 34} < 1$; explanation: Ms = 4.06 and 5.33; $F_{1, 36} = 13.34$, $p < .0008$).

For message source, we conducted a second separate-sample test to determine whether respondents view the sponsored nonprofit (versus the sponsoring firm) as less self-interested in announcing the sponsorship. In a within-subjects design, 60 respondents reported how they would perceive the motivations of Alpo and the Special Olympics as a message source. They rated (on seven-point scales) the likelihood that the source would "have self-serving motives," "be trying to win people's liking or patronage," "be a trustworthy source of information," "make the sponsorship sound better than it is," "have unselfish motives," and "be simply trying to inform the public" (we reverse-scaled some items; Cronbach's $\alpha = .70$). As we expected, respondents viewed the nonprofit as less likely to have self-interested motives (Ms = 3.15 and 4.42 for the Special Olympics and Alpo, respectively; $F_{1, 48} = 44.96$, $p < .0001$). Nonetheless, further investigation into the antecedents of the favorable nonprofit source effect is needed. Study 2's respondents did not note any difference with regard to motives in their undirected thoughts, so it remains possible that the beneficial effect of the nonprofit source is derived purely from affect transfer or some mediating cognition other than thoughts about self-interest. However, that there are beneficial effects is not in question; benefits show up in both scaled measures and undirected thoughts.

Persistence of effects. One year later, we met with 47 of the original 120 respondents, again during professional training seminars. They were evenly spread across conditions. Given this small sample, we examine only effects that are significant on our immediate measures for our whole sample and for the portion that constitutes our follow-up sample. We then use a repeated measures analysis (immediate versus delayed) to track changes in our follow-up sample over time. Otherwise, we might mistakenly conclude that an effect decays over time when, instead, it is absent in our delayed measures because of low power or sample bias. Using this criterion, we are able to examine only effects of fit and cannot examine control-group comparisons.

The impact of fit on firm equity lessens over time ($F_{1, 34} = 5.89$, $p < .02$), but it remains highly significant after a delay of one year. Immediately after exposure to the sponsorship information, this sample's equity scores for Alpo were 4.96 for low fit and 6.39 for created fit ($F_{1, 35} = 35.16$, $p < .0001$). One year later, these equity scores were 5.19 and 5.96 ($F_{1, 34} = 10.52$, $p < .002$).

We also collected sponsorship recall measures. As H_5 predicted, recall is greater with created fit than with low fit. This effect occurs for both free recall and aided recall. Furthermore, the previously documented tendency (Johar and Pham 1999) to incorrectly recall semantically related pairings (i.e., false hits) is reduced by created fit. Free recall is scaled as follows: When a respondent recalled no sponsorship or recalled only that the sponsored cause "had something to do with pets," we code recall as incorrect, or as 0. Although the sponsorship involved the gift of pets and pet food, the Special Olympics is not an organization that "has something to do with pets." When a respondent recalled that the sponsored cause "had something to do with the mentally disabled" or "had something to do with athletics," we scored this response as partially correct, or as 1, because the Special Olympics deals with both. When the respondent recalled that the sponsored cause was the Special Olympics, we scored this response as accurate, or as 2. No one recalled a specific nonprofit other than the Special Olympics. As we predicted, free recall is better with created fit than with low fit (Ms = .60 and .09; $F_{1, 35} = 4.78$, $p < .03$). We reach the same conclusion if we simply compare the proportions of respondents correctly recalling the Special Olympics (35% versus 5%; $\chi^2 = 5.28, p < .02$).

For the aided-recall task, multiple-choice options were "no sponsorship," "the Special Olympics," and "the Humane Society." We coded aided recall as incorrect if respondents indicated that there was no sponsorship or that the sponsored nonprofit was the Humane Society and as correct if they indicated that the sponsored nonprofit was the Special Olympics. The proportion of respondents who recalled correctly is significantly greater with created fit than with low fit (68% and 26%; $\chi^2 = 5.84, p < .01$). Furthermore, the proportion of respondents who incorrectly identified the semantically related Humane Society as the sponsored nonprofit is marginally lower in the created-fit condition than in the low-fit condition (30% and 57%; $\chi^2 = 3.08, p < .07$).

Discussion

Created fit and the persistence of effects. As is the case with high natural fit, created fit increases the favorability of responses relative to low-fit sponsorships and to no sponsorship, at least for familiar, well-liked firms. Furthermore, the created-fit-sponsorship was better recalled than the identical low-fit pairing, consistent with previous work that shows that semantically related pairings are better recalled. When recall is high, we might expect that less communication is required to maintain sponsorship effects over time. As a rough test of this possibility, we examined the extent to which an aided-recall question helped respondents who had not previously free-recalled the sponsorship. Although an aided-recall question helped only 21% of the low-fit respondents, it helped 57% of the created-fit respondents ($\chi^2 = 3.74, p < .05$). Thus, simple communication reminders or point-of-purchase cues may be effective in keeping created-fit (and perhaps high-fit) sponsorships salient. Given the strong effects we observed, further investigation of created fit is warranted. For example, although it seems unlikely, it is possible that the heightened recall resulted from the provision of additional interesting information in the sponsorship announcement rather than from the creation of fit, per se.

Notably, one-year follow-up measures of firm equity indicate that low-fit sponsorships may also have enduring effects, even though they are not as well recalled. Equity means for both low fit and created fit changed little over time. That the effects of fit remained strong over a year during which respondents were likely exposed to other brand messages is impressive.

Message source. Communicating through the sponsored nonprofit can also be beneficial, though it does not completely overcome the negative effects of low fit. The main benefit is a reduction in the impression of misfit and inappropriateness for low-fit sponsorships, as is shown by the perceived fit and clarity measures and by the respondents' undirected thoughts. We also observed benefits of a nonprofit source for high-fit sponsorships, but the absence of these effects in our thought measures leads us to interpret these benefits more cautiously. However, this lack of verbalization may reflect a ceiling effect; high-fit sponsorships are viewed favorably even when the company is the message source.

General Discussion

Much interest in recent years has centered on building, maintaining, and leveraging brand equity. Although much of the brand extension (e.g., Broniarczyk and Alba 1994; John, Loken, and Joiner 1998; Loken and John 1993), brand alliance (e.g., Simonin and Ruth 1998), and cobranding (e.g., Samu, Krishnu, and Smith 1999) literature has focused on brand-specific associations, the corporate social responsibility literature has focused on creating broad associations, such as being "socially responsible," and on the simple transfer of positive affect from sponsored causes (Brown and Dacin 1997; Rifon et al. 2004; Smith 2004; cf. Lichtenstein, Drumwright, and Braig 2004; Sen and Bhattacharya 2001). We adopt a more traditional branding perspective and show that the fit between a firm's specific associations and a sponsored cause can reinforce or blur the firm's positioning. We also show that the unexpectedness of low fit leads to increased elaboration on the sponsorship, that this elaboration is negatively biased, and that it is accompanied by less favorable attitudes toward the sponsorship. Finally, we examine the effects of message content and source and offer the first experimental evidence of the persistence of sponsorship effects over an extended time.

Overview of Results and Managerial Implications

In Study 1, we show that sponsorship fit can blur or reinforce the firm's positioning, influence liking for the sponsorship, and, through these effects, influence firm equity. These results have straightforward implications for sponsorship management. First, fit should be a key consideration in the development of sponsorships. Supporting a familiar, well-liked cause is not enough to ensure a good outcome and may even be harmful. Low fit, even with a well-liked cause, can dilute the firm's positioning, create dislike for the sponsorship, and, as a result, lower firm equity. In contrast, high fit reinforces the firm's positioning, creates favorable attitudes toward the sponsorship, and builds firm equity.

Second, management should pay special attention to factors that influence liking for the sponsorship. Sponsorships influence firm equity through their effects on clarity of positioning and attitude toward the sponsorship. Although high fit reinforces the firm's positioning, it may be only one of many inputs to liking for the sponsorship. For example, choosing a well-liked cause, avoiding unattractive elements (e.g., heavy-handed promotions), and providing engaging ways for consumers to participate (e.g., walkathons, social events) may all contribute to a more favorable attitude toward the sponsorship.

In Study 2, we show that message content and source can mitigate the negative effects of low fit, including poor sponsorship recall, and we demonstrate that sponsorship effects can persist for as long as a year despite day-to-day exposure to other brand communications. Again, managerial implications are straightforward. First, if the firm wishes to sponsor a low-natural-fit cause, management should create fit by making shared associations salient. Created-fit sponsorships can obtain essentially the same benefits as high-natural-fit sponsorships, including good recall, increased clarity of positioning, more favorable attitudes, and greater firm equity. There is plenty of room for creativity here. We created fit between Alpo and the Special Olympics by explaining that caring for pets increases the self-esteem of the mentally disabled. The Ford Motor Company likely created fit with a national park when it explained that it used its expertise in transportation to support the park's transportation needs.

Second, it is risky to assume an understanding of which causes people will perceive as high fit. Instead, it is important to verify how the intended audience will view the relationship. DenTek Oral Care's sponsorship of the American Diabetes Association probably seems logical (i.e., high fit) to people within these two organizations but not to people who are unfamiliar with the dental health effects of diabetes. For this audience, fit needs to be created. The caveat to verify fit also applies to actions that are intended to create fit. We expected that the gift of a pet and pet food to the Special Olympics participants would have a substantial impact on perceived fit with Alpo. The gift did increase perceived fit, but its effect was small compared with that of the explanation that caring for pets increases self-esteem.

Third, firms should consider using the sponsored cause as a message source. This decision may boost the effectiveness of high-fit sponsorships and can reduce some of the risk if the firm is committed to a low-fit sponsorship for which it is difficult to create fit. This strategy may be particularly helpful when transitioning into a high-fit sponsorship portfolio. Finally, it appears that sponsorships do not require extensive ongoing communications support to have enduring effects. We show that sponsorship effects can remain strong for up to one year with no interim reinforcement and as consumers are subject to ordinary day-to-day exposure to brand communications. Particularly for created-fit sponsorships, simple reminders may have strong effects. Free recall of these sponsorships was high, but aided recall, which provided reminders in the form of multiple-choice options, was considerably higher.

Limitations and Research Issues

We examined the effects of one brief exposure to a sponsorship. Given that we find effects on firm equity one year later, the real-world potential for sponsorship effects, both positive and negative, seems great. However, our results can most confidently be generalized to situations such as those we examined: familiar and well-liked firms and causes, causes that do not evoke too many fearful thoughts (e.g., no images of horrible disease), minimal exposures to basic sponsorship information, overall communication levels similar to those for the firms and causes we examined, very-low-fit and very-high-fit sponsorships, and so forth. In particular, we would not expect our predictions and findings with regard to the effect of clarity of positioning to be replicated for firms whose positioning is disliked rather than liked by consumers. Furthermore, although we focused on what we believe are the most common mediators of fit effects—clarity of positioning and attitude toward the sponsorship—it seems likely that other mediators operate in some contexts. Next, we discuss a few of the variables that seem most promising for further research.

Repeated exposures to less simple communications. When a brand extension is launched, it is common for consumers to be repeatedly exposed to well-developed advertisements. In contrast, the typical extension study offers a single exposure to a simple product description. Studies that create a more realistic information environment find that the negative effect of low fit (i.e., low category similarity) decreases as information about the extension, exposures, and consumer innovativeness increase (Klink and Smith 2001; Lane 2000). Although social sponsorships rarely receive communication support equal to that of brand extensions—indeed, minimal exposures are common—repeated and well-produced messages may mitigate the effect of low fit. It is possible to point to apparently successful sponsorships in which fit is questionable, for example, the support of the arts and cultural events by financial institutions. People may have learned through repeated exposure that financial institutions and cultural events go together. If repeated exposure to low-fit relationships can create fit, high levels of promotion may offer another route for overcoming the risks of low fit. However, this strategy may require a level of commitment best reserved for special situations, for example, when a firm seeks to overcome an image problem.

High fit/negative associative links. Some sponsorships are intended to overcome specific negative associations (Hoeffler and Keller 2002; Lichtenstein, Drumwright, and Braig 2004). For example, oil and chemical companies, which are often associated with environmental pollution, frequently sponsor environmental causes. Notably, our pretests indicate that people perceive these kinds of relationships as high fit. For example, manufacturers of alcoholic beverages were rated as having a high fit with Mothers Against Drunk Driving. An important question for future research is whether high fit in these situations has the same beneficial effects we observed in the current studies or whether such sponsorships backfire by reminding consumers of the negative associations. In related findings, Szykman, Bloom, and Blazing (2004) report that respondents infer more society-serving motives when a message about the dangers of drinking and driving is delivered by Mothers Against Drunk Driving than when it is delivered by Budweiser.

The role of persuasion knowledge. In the domains of brand extensions and corporate social responsibility, it has been argued that low fit engenders the attribution that the company's action (e.g., creation of a brand extension, support of a social cause) is a self-interested attempt to gain market value rather than a genuine attempt to deliver value (Aaker 1990; Keller and Aaker 1992; Rifon et al. 2004; cf. Brown and Dacin 1997). This prediction is based on two ideas: (1) that people have "persuasion knowledge" (Friestad and Wright 1994, 1995), or intuitive beliefs about marketers' motives and tactics that they use to interpret marketers' actions, and (2) that low fit somehow activates or shapes the use of this knowledge. For example, Rifon and colleagues (2004) argue that the greater elaboration associated with low fit increases the likelihood that people will retrieve persuasion knowledge and use it to infer that the company is motivated by self-interest (cf. Campbell and Kirmani 2000). Relying on scaled measures of perceived motivation, Rifon and colleagues find the predicted effect. In contrast, our thought listings from both Study 1 and Study 2 show no evidence that this effect occurs spontaneously. Instead, we observe the effect only when we employ scaled measures. Furthermore, these scaled measures do not exhibit discriminant validity. Thus, our results suggest that the predicted attributions were not made spontaneously but rather were constructed on the spot, through attitude priming, to respond to our measures (Feldman and Lynch 1988; Simmons, Bickart, and Lynch 1993). We speculate that sponsorship fit may have the predicted effect on inferred motivations only when motivation is made salient. For example, consumers may be vigilant about marketing efforts of a firm that has recently experienced negative publicity, and therefore they may be quick to think about the firm's motives. If inferred motives vary with fit, a low-fit sponsorship may be a particularly poor way to attempt to rebuild the firm's reputation, whereas a high-fit sponsorship may be a particularly effective way.

We also find some evidence that consumers' use of persuasion knowledge may both increase and decrease the marketer's effectiveness. Our respondents' comments reflected the marketing-savvy attitude reflected in today's popular culture. When fit was low, they chided the marketer: "Alpo will not sell more food." When fit was high, they implicitly applauded the marketer's efforts: "By doing this, they may open more of a market for themselves." This result implies that marketers may impede their effectiveness by any activity that consumers regard as unskilled use of marketing tactics and that they may boost their effectiveness by winning consumers' approval of their skill as marketers. Thus, high-fit sponsorships may help firms achieve marketing objectives *because of* consumers' use of persuasion knowledge.

Appendix A
Study 1 Materials

Alpo

The clipping for the high-fit, sponsorship-absent condition is as follows:

> **Businesses and Organizations in the News**
>
> **P&G** has introduced a home-use version of its fruit and vegetable wash for restaurant and food service operations. The wash is formulated to easily remove dirt and unwanted residues, including up to 93% of wax, 95% of handling residues, and is 98% more effective than water at removing pesticides. The product will be available at groceries and mass merchandisers throughout the U.S.
>
> **The American Heart Association** achieved its goal of increasing access to life-saving automated external defibrillators (AEDs) for air travelers with the new FAA requirement that commercial airlines carry AEDs as part of on-board medical emergency equipment. AEDs are small, easy-to-use devises that can analyze heart rhythms of cardiac arrest victims to determine if shock is necessary and, if warranted, deliver a life-saving jolt of electricity to the heart.
>
> **The Humane Society** today announced the promotion of Elizabeth Hansen to the post of Executive Vice President. Ms. Hansen has been an advocate for the humane treatment of animals for over 20 years. In her five years as Program Director of the Humane Society, she has spearheaded educational campaigns to increase awareness of animal protection issues.
>
> **Alpo Petfoods** today announced the opening of its expanded online store, www.Alpo.com. The store enables consumers to securely purchase an assortment of pet food and related products online. The site includes the ability for returning customers to purchase without re-entering personal information and to interact in real time with customer service representatives.

In low-fit conditions, we replaced the Humane Society item with the following:

> **The Special Olympics** today announced the promotion of Elizabeth Hansen to the post of Executive Vice President. Ms. Hansen has been an advocate for athletic programs for the mentally retarded for over 20 years. In her five years as Program Director of the Special Olympics, she has spearheaded educational campaigns to increase awareness of the benefits of athletics for the mentally retarded.

In sponsorship-present conditions, we added the following to the Alpo item:

> Coincidentally, Alpo also announced its new sponsorship of the [Humane Society or Special Olympics for high-fit versus low-fit conditions], which it initiated with a $25,000 cash donation. The company will also institute a program whereby its employees can receive paid leave to participate as volunteers in [Humane Society, Special Olympics] events.

The Sports Authority

The clipping for the high-fit, sponsorship-absent condition is as follows:

> **Businesses and Organizations in the News**
>
> **P&G** has introduced a home-use version of its fruit and vegetable wash for restaurant and food service operations. The wash is formulated to easily remove dirt and unwanted residues, including up to 93% of wax, 95% of handling residues, and is 98% more effective than water at removing pesticides. The product will be available at groceries and mass merchandisers throughout the U.S.
>
> **The American Heart Association** achieved its goal of increasing access to life-saving automated external defibrillators (AEDs) for air travelers with the new FAA requirement that commercial airlines carry AEDs as part of on-board medical emergency equipment. AEDs are small, easy-to-use devises that can analyze heart rhythms of cardiac arrest victims to determine if shock is necessary

and, if warranted, deliver a life-saving jolt of electricity to the heart.

The Special Olympics today announced the promotion of Elizabeth Hansen to the post of Executive Vice President. Ms. Hansen has been an advocate for athletic programs for the mentally retarded for over 20 years. In her five years as Program Director of the Special Olympics, she has spearheaded educational campaigns to increase awareness of the benefits of athletics for the mentally retarded.

The Sports Authority today announced the opening of its expanded online store, www.TheSportsAuthority.com. The store enables consumers to securely purchase an assortment of sporting goods online. The site includes the ability for returning customers to purchase without re-entering personal information and to interact in real time with customer service.

In low-fit conditions, we replaced the Special Olympics item with the following:

The Humane Society today announced the promotion of Elizabeth Hansen to the post of Executive Vice President. Ms. Hansen has been an advocate for the humane treatment of animals for over 20 years. In her five years as Program Director of the Humane Society, she has spearheaded educational campaigns to increase awareness of animal protection issues.

In sponsorship-present conditions, we added the following to the Sports Authority item:

Coincidentally, The Sports Authority also announced its new sponsorship of the [Special Olympics or Humane Society for high-fit versus low-fit conditions], which it initiated with a $25,000 cash donation. The company will also institute a program whereby its employees can receive paid leave to participate as volunteers in [Special Olympics, Humane Society] events.

Appendix B
Study 2 Materials

The clipping for the no-sponsorship control condition is identical to the low-fit, sponsorship-absent control for Alpo from Study 1:

Businesses and Organizations in the News

P&G has introduced a home-use version of its fruit and vegetable wash for restaurant and food service operations. The wash is formulated to easily remove dirt and unwanted residues, including up to 93% of wax, 95% of handling residues, and is 98% more effective than water at removing pesticides. The product will be available at groceries and mass merchandisers throughout the U.S.

The American Heart Association achieved its goal of increasing access to life-saving automated external defibrillators (AEDs) for air travelers with the new FAA requirement that commercial airlines carry AEDs as part of on-board medical emergency equipment. AEDs are small, easy-to-use devises that can analyze heart rhythms of cardiac arrest victims to determine if shock is necessary and, if warranted, deliver a life-saving jolt of electricity to the heart.

The Special Olympics today announced the promotion of Elizabeth Hansen to the post of Executive Vice President. Ms. Hansen has been an advocate for athletic programs for the mentally retarded for over 20 years. In her five years as Program Director of the Special Olympics, she has spearheaded educational campaigns to increase awareness of the benefits of athletics for the mentally retarded.

Alpo Petfoods today announced the opening of its expanded online store, www.Alpo.com. The store enables consumers to securely purchase an assortment of pet food and related products online. The site includes the ability for returning customers to purchase without re-entering personal information and to interact in real time with customer service representatives.

In the low-fit, company-source condition, the Special Olympics item is unchanged from the control, and we added the following to the Alpo item:

Coincidentally, Alpo also announced its new sponsorship of the Special Olympics, which it initiated with a $25,000 cash donation. The company will also institute a program whereby its employees can receive paid leave to participate as volunteers in Special Olympics events.

In the low-fit, nonprofit-source condition, the Special Olympics item contains the following addition (and the Alpo item is slightly edited):

The Special Olympics also announced Alpo's new sponsorship of the organization, which it initiated with a $25,000 cash donation. The company will also institute a program whereby its employees can receive paid leave to participate as volunteers in Special Olympics events.

Coincidentally, **Alpo Petfoods** today announced the opening of its expanded online store, www.Alpo.com. The store enables consumers to securely purchase an assortment of pet food and related products online. The site includes the ability for returning customers to purchase without re-entering personal information and to interact in real time with customer service representatives.

In the created-fit, company-source condition, the Special Olympics item is again unchanged from the control, and we added the following to the Alpo item:

Coincidentally, Alpo also announced its new sponsorship of the Special Olympics, which it initiated with a $25,000 cash donation. Citing research that shows that caring for pets increases the self-confidence of the mentally disabled, Alpo reported that it will arrange for interested Special Olympics participants to receive a dog or cat and a lifetime supply of pet food. The company will also institute a program whereby its employees can receive paid leave to participate as volunteers in Special Olympics events.

In the created-fit, nonprofit-source condition, the Special Olympics item contains the following addition (and the Alpo item is slightly edited as previously):

The Special Olympics also announced Alpo's new sponsorship of the organization, which it initiated with a $25,000 cash donation. Citing research that shows that caring for pets increases the self-confidence of the mentally disabled, the organization reported that Alpo will arrange for interested Special Olympics participants to receive a dog or cat and a lifetime supply of pet food. The company will also institute a program whereby its employees can receive paid leave to participate as volunteers in Special Olympics events.

Coincidentally, **Alpo Petfoods** today announced the opening of its expanded online store, www.Alpo.com. The store enables consumers to securely purchase an assortment of pet food and related products online. The site includes the ability for returning customers to purchase without re-entering personal information and to interact in real time with customer service representatives.

REFERENCES

Aaker, David A. (1990), "Brand Extensions: The Good, the Bad, and the Ugly," *Sloan Management Review*, 31 (Summer), 47–56.

——— and Kevin Lane Keller (1990), "Consumer Evaluations of Brand Extensions," *Journal of Marketing*, 54 (January), 27–41.

Anderson, James C. and David W. Gerbing (1988), "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin*, 103 (May), 411–23.

Artz, Nancy and Alice M. Tybout (1999), "The Moderating Impact of Quantitative Information on the Relationship Between Source Credibility and Persuasion: A Persuasion Knowledge Model Interpretation," *Marketing Letters*, 10 (February), 51–62.

Bentler, Peter M. (1989), *EQS Structural Equations Program*. Los Angeles: BMDP Statistical Software.

——— and Douglas G. Bonett (1980), "Significance Tests and Goodness-of-Fit in the Analysis of Covariance Structures," *Psychological Bulletin*, 88 (November), 588–606.

Birnbaum, Michael H. and Steven E. Stegner (1979), "Source Credibility in Social Judgment: Bias, Expertise, and the Judge's Point of View," *Journal of Personality and Social Psychology*, 37 (January), 48–74.

Boush, David M. and Barbara Loken (1991), "A Process-Tracing Study of Brand Extension Evaluation," *Journal of Marketing Research*, 28 (February), 16–28.

Bridges, Sheri, Kevin Lane Keller, and Sanjay Sood (2000), "Communication Strategies for Brand Extensions: Enhancing Perceived Fit by Establishing Explanatory Links," *Journal of Advertising*, 24 (Winter), 1–11.

Broniarczyk, Susan M. and Joseph W. Alba (1994), "The Importance of the Brand in Brand Extension," *Journal of Marketing Research*, 31 (May), 214–28.

Brown, Tom J. and Peter A. Dacin (1997), "The Company and the Product: Corporate Associations and Consumer Product Responses," *Journal of Marketing*, 61 (January), 68–84.

Campbell, Margaret C. and Amna Kirmani (2000), "Consumer's Use of Persuasion Knowledge: The Effects of Accessibility and Cognitive Capacity on Perceptions of an Influence Agent," *Journal of Consumer Research*, 27 (June), 69–83.

DenTek Oral Care (2006), "A Proud Sponsor of the ADA," (accessed March 19, 2006), [available at http://www.usdentek.com/home_ada.htm#dentek]

Doty, D. Harold and William H. Glick (1998), "Common Methods Bias: Does Common Methods Variance Really Bias Results?" *Organizational Research Methods*, 1 (October), 374–406.

Erdem, Tulin (1998), "An Empirical Analysis of Umbrella Branding," *Journal of Marketing Research*, 35 (August), 339–51.

——— and Joffre Swait (1998), "Brand Equity as a Signaling Phenomenon," *Journal of Consumer Psychology*, 7 (2), 131–57.

Fazio, Russell H. and Carol J. Williams (1986), "Attitude Accessibility as a Moderator of the Attitude-Perception and Attitude-Behavior Relations: An Investigation of the 1984 Presidential Election," *Journal of Personality and Social Psychology*, 51 (September), 505–514.

Feldman, Jack M. and John G. Lynch Jr. (1988), "Self-Generated Validity and Other Effects of Measurement on Belief, Attitude, Intention, and Behavior," *Journal of Applied Psychology*, 73 (August), 421–35.

Ford Motor Company (2006), "Ford Technology Puts American Icon on the Road Again," (May 3, 2002), (accessed March 19, 2006), [available at http://media.ford.com/newsroom/release_display.cfm?release=12150].

Fornell, Claes and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39–50.

Friestad, Marian and Peter Wright (1994), "The Persuasion Knowledge Model: How People Cope with Persuasion Attempts," *Journal of Consumer Research*, 21 (June), 1–31.

——— and ——— (1995), "Persuasion Knowledge: Lay People's and Researcher's Beliefs About the Psychology of Advertising," *Journal of Consumer Research*, 22 (June), 62–74.

Gilbert, Daniel T. (1989), "Thinking Lightly About Others: Automatic Components of the Social Inference Process," in *Unintended Thought*, James S. Uleman and John A. Bargh, eds. New York: Guilford Press, 189–211.

Hatcher, Larry (1994), *A Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling*. Cary, NC: SAS Institute Inc.

Hoeffler, Steve and Kevin Lane Keller (2002), "Building Brand Equity Through Corporate Societal Marketing," *Journal of Public Policy & Marketing*, 21 (Spring), 78–89.

Houston, David A. and Russell H. Fazio (1989), "Biased Processing as a Function of Attitude Accessibility: Making Objective Statements Subjectively," *Social Cognition*, 7 (Spring), 51–66.

Jagre, Emma, John J. Watson, and John G. Watson (2001), "Sponsorship and Congruity Theory: A Theoretical Framework for Explaining Consumer Attitude and Recall of Event Sponsorship," in *Advances in Consumer Research*, Vol. 28, Mary C. Gilly and Joan Meyers-Levy, eds. Valdosta, GA: Association for Consumer Research, 439–45.

Johar, Gita Venkataramani and Michel Tuan Pham (1999), "Relatedness, Prominence, and Constructive Sponsor Identification," *Journal of Marketing Research*, 36 (August), 299–312.

——— and Carolyn J. Simmons (2000), "The Use of Concurrent Disclosures to Correct Invalid Inferences," *Journal of Consumer Research*, 26 (March), 307–322.

John, Deborah Roedder, Barbara Loken, and Christopher Joiner (1998), "The Negative Impact of Extensions: Can Flagship Products Be Diluted?" *Journal of Marketing*, 62 (January), 19–31.

Keller, Kevin Lane (1993), "Conceptualizing, Measuring, and Managing Consumer-Based Brand Equity," *Journal of Marketing*, 57 (January), 1–10.

——— and David A. Aaker (1992), "The Effects of Sequential Introduction of Brand Extensions," *Journal of Marketing Research*, 24 (February), 35–50.

Kenny, David A., Deborah A. Kashy, and Niall Bolger (1998), "Data Analysis in Social Psychology," in *The Handbook of Social Psychology*, Vol. 1, 4th ed., Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey, eds. New York: McGraw-Hill, 233–65.

Kishton, Joseph M. and Keith F. Widaman (1994), "Unidimensional Versus Domain Representative Parceling of Questionnaire Items: An Empirical Example," *Educational and Psychological Measurement*, 54 (Fall), 565–75.

Klink, Richard R. and Daniel C. Smith (2001), "Threats to the External Validity of Brand Extension Research," *Journal of Marketing Research*, 38 (August), 326–35.

Lane, Vicki R. (2000), "The Impact of Ad Repetition and Ad Content on Consumer Perceptions of Incongruent Extensions," *Journal of Marketing*, 64 (April), 80–91.

Lichtenstein, Donald R., Minette E. Drumwright, and Bridgette M. Braig (2004), "The Effect of Corporate Social Responsibility on Customer Donations to Corporate-Supported Nonprofits," *Journal of Marketing*, 68 (October), 16–32.

Loken, Barbara and Deborah Roedder John (1993), "Diluting Brand Beliefs: When Do Brand Extensions Have a Negative Impact?" *Journal of Marketing*, 57 (July), 71–84.

Mandler, George (1982), "The Structure of Value: Accounting for Taste," in *Affect and Cognition: The 17th Annual Carnegie Symposium*, Margaret S. Clark and Susan T. Fiske, eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 3–36.

Marsh, Herbert W., Kit-tai Hau, John R. Balla, and David Grayson (1998), "Is More Ever Too Much? The Number of Indicators per Factor in Confirmatory Factor Analysis," *Multivariate Behavioral Research*, 33 (2), 181–220.

Milberg, Sandra J., C. Whan Park, and Michael S. McCarthy (1997), "Managing Negative Feedback Effects Associated with Brand Extensions: The Impact of Alternative Branding Strategies," *Journal of Consumer Psychology*, 6 (2), 119–40.

Park, C. Whan, Bernard J. Jaworski, and Deborah J. MacInnis (1986), "Strategic Brand Concept-Image Management," *Journal of Marketing*, 50 (October), 135–45.

———, Sandra Milberg, and Robert Lawson (1991), "Evaluation of Brand Extensions: The Role of Product Feature Similarity and Brand Concept Consistency," *Journal of Consumer Research*, 18 (September), 185–93.

Pham, Michael Tuan and Gita Venkataramani Johar (2001), "Market Prominence Biases in Sponsor Identification: Processes and Consequentiality," *Psychology and Marketing*, 18 (February), 123–43.

Porter, Michael E. (1987), "From Competitive Advantage to Corporate Strategy," *Harvard Business Review*, 65 (May–June), 43–51.

Pracejus, John W. and G. Douglas Olsen (2004), "The Role of Brand/Cause Fit in the Effectiveness of Cause-Related Marketing Campaigns," *Journal of Business Research*, 57 (June), 635–40.

Reilly, Jeanne (2000), "Charitable Work Sells at a Number of Firms," *Marketing News*, (September 11), 46.

Rifon, Nora J., Sejung Marina Choi, Carrie S. Trimble, and Hairong Li (2004), "Congruence Effects in Sponsorship: The Mediating Role of Sponsor Credibility and Consumer Attributions of Sponsor Motive," *Journal of Advertising*, 33 (Spring), 29–42.

Rumelt, Richard P. (1974), *Strategy, Structure, and Economic Performance*. Cambridge, MA: Harvard University Press.

Samu, Sridhar, H. Shanker Krishnan, and Robert E. Smith (1999), "Using Advertising Alliances for New Product Introduction: Interactions Between Product Complementarity and Promotional Strategies," *Journal of Marketing*, 63 (January), 57–74.

Sen, Sankar and C.B. Bhattacharya (2001), "Does Doing Good Always Lead to Doing Better? Consumer Reactions to Corporate Social Responsibility," *Journal of Marketing Research*, 38 (May), 225–38.

Simmons, Carolyn J., Barbara A. Bickart, and John G. Lynch (1993), "Capturing and Creating Public Opinion in Survey Research," *Journal of Consumer Research*, 20 (September), 316–29.

Simonin, Bernard and Julie A. Ruth (1998), "Is a Company Known by the Company It Keeps? Assessing the Spillover Effects of Brand Alliances on Consumer Attitudes," *Journal of Marketing Research*, 35 (February), 30–42.

Smith, Gareth (2004), "Brand Image Transfer Through Sponsorship: A Consumer Learning Perspective," *Journal of Marketing Management*, 20 (April), 457–74.

Speed, Richard and Peter Thompson (2000), "Determinants of Sports Sponsorship Response," *Journal of the Academy of Marketing Science*, 28 (Spring) 226–38.

Sullivan, Mary (1990), "Measuring Image Spillovers in Umbrella-Branded Products," *Journal of Business*, 63 (July), 309–329.

Szykman, Lisa R., Paul N. Bloom, and Jennifer Blazing (2004), "Does Corporate Sponsorship of a Socially-Oriented Message Make a Difference? An Investigation of the Effects of Sponsorship Identity on Responses to an Anti-Drinking and Driving Message," *Journal of Consumer Psychology*, 14 (1–2), 13–20.

Weiner, Bernard (1985), "'Spontaneous' Causal Thinking," *Psychological Bulletin*, 97 (January), 74–84.

Weiner, Joshua L., Raymond W. LaForge, and Jerry R. Goolsby (1990), "Personal Communication in Marketing: An Examination of Self-Interest Contingency Relationships," *Journal of Marketing Research*, 27 (May), 227–31.

Rajagopal Raghunathan, Rebecca Walker Naylor, & Wayne D. Hoyer

The Unhealthy = Tasty Intuition and Its Effects on Taste Inferences, Enjoyment, and Choice of Food Products

Across four experiments, the authors find that when information pertaining to the assessment of the healthiness of food items is provided, the less healthy the item is portrayed to be, (1) the better is its inferred taste, (2) the more it is enjoyed during actual consumption, and (3) the greater is the preference for it in choice tasks when a hedonic goal is more (versus less) salient. The authors obtain these effects both among consumers who report that they believe that healthiness and tastiness are negatively correlated and, to a lesser degree, among those who do not report such a belief. The authors also provide evidence that the association between the concepts of "unhealthy" and "tasty" operates at an implicit level. The authors discuss possibilities for controlling the effect of the unhealthy = tasty intuition (and its potential for causing negative health consequences), including controlling the volume of unhealthy but tasty food eaten, changing unhealthy foods to make them less unhealthy but still tasty, and providing consumers with better information about what constitutes "healthy."

Assessing the health consequences of food consumption habits has emerged as an important topic of inquiry in recent years (Oakes 2005; Oakes and Slotterback 2005; Seiders and Petty 2004; Verbeke 2006; Wansink 2004a, b; Wansink and Huckabee 2005). From an evolutionary perspective, the easy availability of food presents a unique quandary for developed countries and societies. On the one hand, it has virtually obliterated problems resulting from malnourishment. On the other hand, it has generated problems stemming from the overconsumption of food. For example, obesity is growing at an unprecedented rate, not just in the United States, where nearly two-thirds of the population are overweight or obese, but in other developing nations as well (Seiders and Petty 2004).

To many consumers and consumer activists, organizations (e.g., lobbyists for grocery manufacturers) and restaurants (e.g., McDonald's, Hardee's) that are interested in promoting unhealthy eating habits are to blame for consumers' propensity to overconsume (Brownell and Horgen 2003; Nestle 2003). Consequently, lawsuits have been filed against fast-food companies such as McDonald's (Parloff 2003). Attorneys representing these companies have argued that consumers are aware of the ingredients in their offerings and that they understand the weight and health consequences of overconsumption. In their view, consumers are ultimately responsible for the negative consequences of their own actions.[1]

The objective of this research is to examine the role of a subtle (and currently unrecognized) factor in the overconsumption of food that is perceived as unhealthy. What if people consume food that is considered unhealthy not despite its perceived unhealthiness but rather *because* of it? That is, what if part of the attractiveness of food lies in its perceived unhealthiness? This can happen if consumers intuitively believe that the unhealthier the food, the tastier it is. The operation of such a belief would increase the chances that people will overconsume food portrayed as less (versus more) healthy because they expect that such food will taste better. Consistent with this idea, we propose that the perceived unhealthiness of food has the ironic effect of enhancing its attractiveness. Consequently, under circumstances in which a hedonic (e.g., enjoyment) goal is more (versus less) salient, people will choose options they perceived as unhealthier, even if no information about their tastiness relative to other options is available. Across the experiments we conduct in this research, we obtain support for this proposition.

Rajagopal Raghunathan is Assistant Professor of Marketing (e-mail: Raj.Raghunathan@mccombs.utexas.edu), and Wayne D. Hoyer is The James L. Bayless/William S. Farish Fund Chair for Free Enterprises and Professor of Marketing (e-mail: wayne.hoyer@mccombs.utexas.edu), McCombs School of Business, University of Texas at Austin. Rebecca Walker Naylor is Assistant Professor of Marketing, Moore School of Business, University of South Carolina (e-mail: rnaylor@moore.sc.edu). The authors thank Susan Broniarczyk, Kristin Diehl, Julie Irwin, Andrew Gershoff, Jonathan Levav, and attendees at the 2005 Marketing Young Scholars Conference for their helpful suggestions on the article.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

---

[1] The fast-food chains appear to be winning the battle, at least for now. In October 2005, the U.S. House of Representatives passed a bill to block lawsuits blaming restaurants and food companies for causing obesity or weight gain (Zhang 2005).

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 70 (October 2006), 170–184

Our research generates two theoretically interesting and substantively meaningful insights. First, from a theoretical standpoint, we document both direct and indirect support for our main thesis, namely, that people subscribe to the unhealthy = tasty intuition and that this intuition affects inferences about food choice, actual enjoyment of foods, and choice of food products. These notions are particularly intriguing because there is little rigorous evidence that healthiness and tastiness are indeed inversely related to each other. Second, our research suggests that the influence of the unhealthy = tasty intuition can be implicit; as such, even people who may not report believing that unhealthy food is tastier make judgments and choices as if they subscribed to such a view.

We structure the rest of the article as follows: We begin by briefly commenting on the potential sources for the unhealthy = tasty intuition. Then, we derive predictions of how the intuition affects three variables of theoretical and substantive interest: (1) taste inferences of food products when information pertaining to their healthiness is provided, (2) differences in enjoyment of the same food product when it is presented as being more versus less healthy, and (3) choice between two snack options as a function of differences in the propensity to pursue a hedonic goal. We report results from four controlled experiments that test these hypotheses and conclude with a discussion of the implications of our findings.

Theory

Sources of the Unhealthy = Tasty Intuition

Intuitions or lay beliefs and theories may be generated internally, through personal experience and self-observation (Ross and Nisbett 1991), or externally, from environmental cues (Morris, Menon, and Ames 2001). We posit that both types of sources underlie the unhealthy = tasty intuition.

Internal source. We posit that the unhealthy = tasty intuition is generated internally on the basis of a belief in a compensatory relationship between the "wholesomeness" of stimuli and their "hedonic potential." That is, we believe that the unhealthy = tasty intuition is a specific manifestation of a more general principle, according to which there is an inverse relationship between things that are "healthy," "nourishing," and "good for you" and those that are "enjoyable," "fun," and "exciting." Evidence for an assumed compensatory relationship between the wholesomeness and hedonic potential of stimuli emerged across a separate series of experiments we conducted. For example, the results from one such experiment revealed that compared with an automobile portrayed as attractive and fun to drive, one that was portrayed as unattractive and less enjoyable to drive was inferred to be safer, even though the more attractive automobile was objectively safer (as per the Insurance Institute of Highway Safety). We speculate that a belief in the more general unwholesome = fun intuition is rooted in religious messages, such as the Protestant work ethic, according to which a person is morally obligated to accord greater priority to necessities over luxuries (e.g., Berry 1994; Weber 1998; see also Maslow 1962). Implicit in this view are the ideas that comfort and luxury do not come easily and that a person needs to work hard to earn the right to indulge (e.g., Kivetz and Simonson 2002). Thus, the Protestant work ethic makes a clear distinction between stimuli or activities that are of a "serious," "worthy," and "work-oriented" nature and those that are of a "trivial," "frivolous," and "fun-related" nature and suggests that stimuli/activities fall neatly into one or the other category but rarely into both. In the context of food, this translates into a belief that healthy food is less tasty.

External source. Operating in concert with the internal source for the unhealthy = tasty intuition is an external one. Through the mass media or personal communication, people are constantly exposed to views that are compatible with the intuition. Several recent articles in popular magazines (e.g., "Is Fat the Next Tobacco?" which appeared in *Fortune* [Parloff 2003]), popular books (e.g., *Fast Food Nation*, by Eric Schlosser [2001]), and movies (e.g., Morgan Spurlock's *Super Size Me*, which appeared in theaters nationwide in 2004) explicitly or implicitly suggest an inverse relationship between healthiness and tastiness. Consumers are also likely to have been exposed to the unhealthy = tasty view through personal communication. For example, parents and guardians appear to find value in propagating such a message because it encourages children to consume hedonically unpleasant food that is perceived as salubrious (e.g., eating broccoli) and discourages them from overconsuming hedonically appealing fare that is potentially harmful (e.g., candy).

In summary, we propose that the unhealthy = tasty intuition is generated internally through its consistency with the more general unwholesome = fun intuition and/or externally through repeated exposure to views that are compatible with it.[2] We believe that these sources help generate the hypothesis that unhealthiness is inversely related to tastiness. After this association has been generated, psychological mechanisms help perpetuate it, as we discuss next.

Influence of the Unhealthy = Tasty Intuition on Decision Making

Figure 1 depicts the influence of the unhealthy = tasty intuition on three conceptually separate stages of the decision-making process. Research has shown that consumers generate inferences about missing product attributes on the basis of lay theories, that is, beliefs and intuitions about the world (Broniarczyk and Alba 1994; Sujan and Dekleva 1987). Indeed, inferences about missing product attributes are driven by intuitions even in the face of contradictory cues. For example, in a study in which participants were asked to make inferences about a camera's durability, the participants relied more heavily on the intuition that durability and warranty are positively correlated than on the actual correlation (in the presented data) between durability and shutter speed, an association that is not supported by intuition (Broniarczyk and Alba 1994).

---

[2] It is worthwhile to mention an alternative source for the intuition: direct experience. A person may arrive at the intuition through lay inferencing (Kelly and Thibaut 1969), a process that involves personal observation that unhealthier food is tastier. There are two reasons we do not believe that this is a valid source for the intuition. First, lay inferencing presupposes objectivity on the part of the inference maker. However, we know that because of the prevalence of external messages consistent with the unhealthy = tasty intuition, consumers are unlikely to be objective judges of the correlation between healthiness and tastiness. Second, to assess accurately the true correlation between healthiness and tastiness, it is first necessary to assign relatively stable healthiness and tastiness ratings to various types of food. However, relevant findings indicate that ratings of the healthiness and tastiness of food likely depend heavily on contextual factors (Drewnowski 1997; Oakes 2005).

Influence of intuition on taste inferences. Consistent with these findings, we predict that as a consequence of belief in the unhealthy = tasty intuition (see Figure 1, Stage 1),

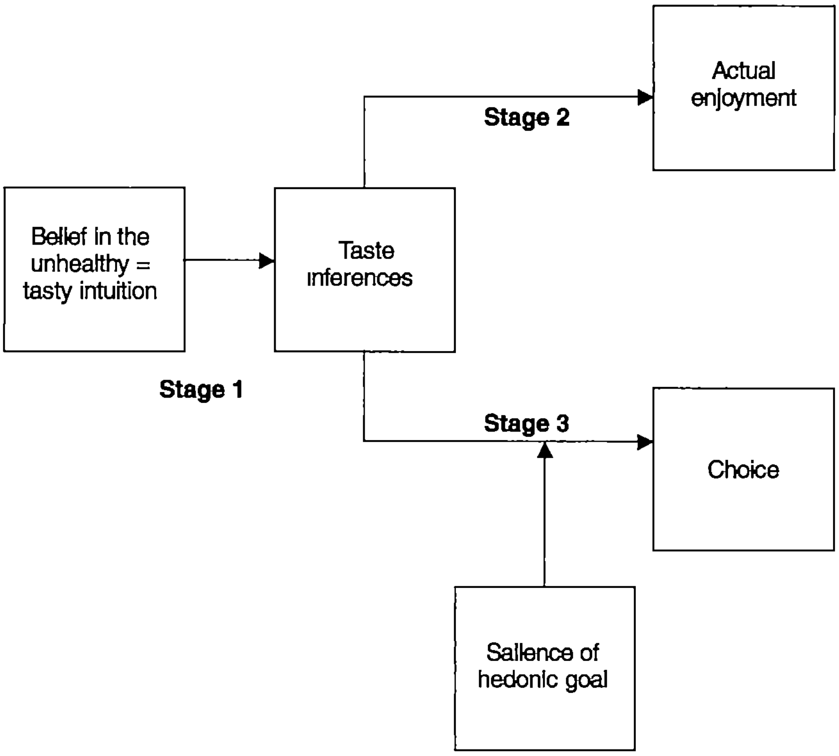
> H_1: The healthier a food item is perceived to be, the lower is its inferred tastiness.

Influence of intuition on enjoyment. Inferences about missing attributes are likely to act as hypotheses that can bias judgments about these attributes (Deighton 1984; Hoch and Ha 1986; Shiv, Carmon, and Ariely 2005). The operation of this hypothesis confirmation bias was captured in Shiv, Carmon, and Ariely's (2005) experiments, which were designed to demonstrate the placebo effects of marketing actions. For example, in their first experiment, they find that consumers who pay a discounted price for an energy drink that purportedly increases mental acuity derive less benefit from the drink (measured in terms of number of mental puzzles solved) than do participants who pay full price. Consistent with the operation of the hypothesis confirmation bias, the discounted price serves as a hypothesis about the (lack of) effectiveness of the drink, which is confirmed on consumption of the drink (and reflected in the subsequent puzzle-solving task). In the context of the current research, we expect that consumers will judge the same food product as more tasty when it is portrayed as less healthy. We state this prediction more formally as follows (see Figure 1, Stage 2):

> H_2: When actual tastiness of a food product is controlled for, the healthier the food is portrayed to be, the lower is its judged tastiness.

Influence of intuition on choice. Product choice is often a function of consumption goals (Bettman, Luce, and Payne 1998). In the context of food choice, researchers have suggested that, in general, the hedonic goal of enjoyment is salient and important (Dhar and Simonson 1999; Food Marketing Institute 1996; Shiv and Fedorikhin 1999; Tepper and Trail 1998). For example, Shiv and Fedorikhin's (1999) experiments examine how a person's mood can influence how he or she resolves the trade-off between the enjoyment and the health benefits of food offerings. Likewise, Dhar and Simonson's (1999) experiments examine conditions that lead to "highlighting" (a focus on either an enjoyment

FIGURE 1
Model of the Influence of the Unhealthy = Tasty Intuition on Judgments and Decisions



or a health goal in choosing an entrée and a dessert) versus "balancing" (a focus on both goals). Given that enjoyment is likely to be a significant goal when people choose food products, the unhealthy = tasty intuition suggests that consumers with a more (versus less) salient hedonic goal will have a relatively greater preference for options they perceived as unhealthy (see Figure 1, Stage 3).

> H_3: Given a choice between two food products that vary in healthiness, people with a more (versus less) salient enjoyment goal will choose the less healthy option.

Implicit and explicit influence of intuition. Thus far, we have suggested that consumers' taste inferences, the judged tastiness of food, and product choice will be influenced in a manner consistent with the unhealthy = tasty intuition. Here, we explore whether explicitness of belief in the intuition moderates this general pattern. If we find that even consumers who do not explicitly report a belief in the intuition make decisions as if they did, important theoretical and substantive implications would emerge. From a theoretical standpoint, such a finding would suggest that the influence of the intuition can occur outside the awareness of consumers making consumption decisions in the context of food. From a substantive standpoint, the finding would point to the difficulty of controlling the influence of the intuition.

The possibility that even people who do not report a belief in the intuition will make decisions as if they believed in it is consistent with the mechanism of hidden covariation detection (e.g., Lewicki 1986; Lewicki and Hill 1989; Lewicki, Hill, and Czyzewska 1997). Lewicki and colleagues find that people can learn the relationship between variables implicitly, that is, without being able to articulate their learning. In one experiment (Lewicki 1986, Experiment 1a), participants were exposed to information about some traits (e.g., kindness, capability) of women such that these traits covaried with their hair length. Later, the same participants were asked to assess these traits in another set of women with long and short hair. Participants' judgments of these targets were consistent with the covariations between hair lengths and traits in the initial stimuli. Importantly, even participants who were unable to articulate these covariations explicitly made judgments as if they had learned them. Thus, people are able to acquire information nonconsciously about covariations in a data set without being able to report such learning explicitly (Lewicki and Hill 1989, p. 240).

Consistent with Lewicki and colleagues' (e.g., Lewicki 1986; Lewicki and Hill 1989; Lewicki, Hill, and Czyzewska 1997) work, we expect that as a result of exposure to the many external sources that suggest a negative correlation between healthiness and tastiness, consumers will have implicitly learned to associate unhealthiness with tastiness. Thus, we expect that even consumers who do not explicitly report believing that unhealthier food is tastier will make judgments and decisions as if they did. However, because there is evidence that explicitly (versus implicitly) learned and articulated relationships are likely to have significantly stronger influences on judgments (Allen and Janiszewski 1989), we expect that the influence of the unhealthy = tasty intuition will be stronger among people who report explicit belief in it. Thus:

> H_4: The pattern of results we predict in H_1–H_3 holds for both people who report an explicit belief in the unhealthy = tasty intuition and those who do not, but it is stronger among the former group.

Experiment 1: Implicitness of Belief in Intuition

We designed Experiment 1 to establish that people implicitly subscribe to the unhealthy = tasty intuition. To address this objective, we used the implicit association test (IAT; Greenwald, McGhee, and Schwartz 1998; Greenwald, Nosek, and Banaji 2003), an established tool for demonstrating the type of implicit association we wish to demonstrate.

Participants and Procedure

One hundred thirty-eight undergraduate students participated in this experiment (which we conducted using IBM computers and the Direct RT software) for course extra credit. Participants were told that their task was to categorize correctly stimuli shown on the middle of their screens. We selected stimuli from one of the following four categories (see the Appendix) for the categorization tasks: (1) pictures of unhealthy foods, (2) pictures of healthy foods, (3) words associated with the tastiness or enjoyment of food (e.g., tasty, delicious), and (4) words associated with the lack of taste or enjoyment (e.g., flavorless, unpalatable). In line with established protocol (Greenwald, McGhee, and Schwartz 1998), participants completed seven blocks of trials, five of which were practice blocks designed to familiarize participants with the target stimuli and the categorization labels. The critical blocks for analysis were Block 4, in which the category labels were "Healthy Foods/Enjoyable" versus "Unhealthy Foods/Not Enjoyable," and Block 7, in which the category labels were "Unhealthy Foods/ Enjoyable" and "Healthy Foods/Not Enjoyable" (see Table 1).

Stimuli from all four categories (healthy food images, unhealthy food images, enjoyment words, and lack of enjoyment words) were presented for classification in the critical blocks for 48 trials across both blocks. An implicit belief in the intuition that unhealthy foods are tastier (and that healthier foods are less tasty) would be reflected in faster responses when stimuli were to be categorized in the condition in which the classification task was congruent with the intuition (i.e., when participants paired unhealthy foods with words associated with tastiness and enjoyment) than in the incongruent condition (i.e., when participants paired healthy foods with words associated with tastiness and enjoyment). Explicit measures of belief in the unhealthy = tasty intuition were collected (using paper-and-pencil measures) before the IAT procedure to compare participants' explicit and implicit attitudes. We assessed explicitness of participants' beliefs in the intuition by eliciting their agreement with the following two items (1 = "strongly disagree," 9 = "strongly agree"): (1) "Things that are good

TABLE 1
Experiment 1: Sequence of Trial Blocks in IAT

Block	Number of Trials	Function	Items Assigned to Left-Key Response	Items Assigned to Right-Key Response
1	20	Practice	Healthy food images	Unhealthy food images
2	28	Practice	Good taste/enjoyment words	Bad taste/lack of enjoyment words
3	24	Practice	Healthy food images + good taste/enjoyment words	Unhealthy food images + bad taste/lack of enjoyment words
4	48	Critical test block	Healthy food images + good taste/enjoyment words	Unhealthy food images + bad taste/lack of enjoyment words
5	40	Practice	Unhealthy food images	Healthy food images
6	24	Practice	Unhealthy food images + good taste/enjoyment words	Healthy food images + bad taste/lack of enjoyment words
7	48	Critical test block	Unhealthy food images + good taste/enjoyment words	Healthy food images + bad taste/lack of enjoyment words

for me rarely taste good," and (2) "There is no way to make food healthier without sacrificing taste."

Data Preparation

Following the revised IAT scoring algorithm (Greenwald, Nosek, and Banaji 2003), we subjected all data to the following criteria: (1) We eliminated trial response latencies greater than 10,000 milliseconds (of the 13,248 trials in the critical blocks, 2 were above this threshold and thus were eliminated); (2) we excluded participants whose response times were less than 300 milliseconds on more than 10% of the critical trials (we discarded 2 of the participants, resulting in a final sample of 136 participants); (3) we included all response latencies, even those for false responses; and (4) we computed a difference score between the two critical blocks of trials ([congruent: unhealthy food + enjoyable and healthy foods + not enjoyable] – [incongruent: healthy foods + enjoyable and unhealthy foods + not enjoyable] divided by the pooled standard deviation of response latencies across both blocks). The resultant measure is the IAT D effect (Greenwald, Nosek, and Banaji 2003).

Results and Discussion

Response latencies were significantly lower when the stimuli were presented in the congruent (versus incongruent) condition. Mean response time when participants were asked to classify stimuli into the congruent categories (pairing unhealthiness with enjoyment) was 767.76 milliseconds, compared with 1137.91 milliseconds when participants were asked to classify stimuli into the incongruent categories (pairing healthiness with enjoyment); this difference in response time was significant ($D = .648$, $SD = .497$, $t(135) = 17.19$, $p < .0001$, $d = 1.42$). That participants were able to respond significantly faster when asked to group unhealthy foods with words describing tastiness and enjoyment than with words describing lack of taste and lack of enjoyment suggests that participants have a stronger implicit association between unhealthiness and tastiness than they do between healthiness and tastiness.

To test whether the unhealthy = tasty intuition influences the response times of even the participants who did not report explicit belief in the intuition, we first averaged the items used to measure explicitness of belief to form an explicitness-of-belief index. We then assessed the mean difference in response time for the congruent versus incongruent condition at each level of reported explicitness of belief in the intuition (see Irwin and McClelland 2001). The results revealed that this difference was significant at every level of belief. That is, at all reported levels of explicit disagreement and agreement with the intuition, the difference in response latencies when participants were asked to pair healthiness with enjoyment versus when they were asked to pair unhealthiness with enjoyment was significantly different from zero (all t's > 3.23, all $ps < .01$). Participants at all levels of belief were able to classify unhealthiness with enjoyment more quickly than they were able to classify healthiness with enjoyment.

Thus, the results of Experiment 1 are consistent with our hypothesis that consumers implicitly subscribe to the view that healthiness and tastiness are inversely related to each other, regardless of their level of explicit belief in this relationship. As such, we expect that even consumers who do not report agreeing that healthiness and tastiness are inversely related to each other will make judgments and decisions as if they did; we tested this hypothesis in the next three experiments.

Experiment 2: Taste Inferences of Crackers

The results from Experiment 1 demonstrated that people implicitly associate unhealthiness with tastiness. Thus, we expect that the influence of the unhealthy = tasty intuition on decision making can take place outside of awareness such that even participants who do not report believing that healthiness and tastiness are inversely related to each other make decisions as if they believe in such a relationship. In general, the objective of Experiment 2 was to test this prediction in the context of inference making. Specifically, the objective of this experiment was twofold: First, we wanted to assess whether a food product portrayed as unhealthy would be inferred to taste better (H_1) than one portrayed as healthy. Second, we wanted to assess whether such an inference is made both by people who explicitly report believing that tastiness is inversely related to healthiness and, to a lesser degree, by those who do not (H_4).

Participants, Stimuli, and Procedure

One hundred ten undergraduate students (51 men and 59 women) took part in this experiment for course extra credit. Participants were told that the purpose of the experiment was to assess consumers' taste perceptions of three new brands of cheddar-flavored snack crackers. To lend credibility to the cover story, participants were told that a national cracker manufacturer was interested in introducing these brands of crackers into the local market. The brands were portrayed as varying in terms of unhealthiness. After being exposed to nutritional information for the cracker versions, participants were asked to infer the tastiness of the three versions.

Unhealthiness manipulation. The most natural way to manipulate unhealthiness perceptions is to portray offerings as differing in the content of an ingredient (e.g., fat) that is generally accepted as being unhealthy. However, altering the content of fat (or any other ingredient) can alter taste perceptions directly (i.e., without altering perceptions of unhealthiness). As a result, varying content of a specific ingredient would confound our prediction with the competing possibility that the ingredient may independently enhance inferred tastiness of the crackers. To overcome this problem, we chose to manipulate perceptions of unhealthiness by varying the purported content of "bad" versus "good" fat, while controlling for the crackers' total fat content. This allowed us to manipulate the perceived unhealthiness of the crackers selectively while controlling for any (potential) effects due to a direct association of fat content with taste.

The unhealthiness manipulation unfolded as follows: Before seeing the nutritional information for each cracker version, participants were presented with the following information about the characteristics of bad (saturated) and good (unsaturated) fat:

> Medical opinion suggests that consuming unsaturated fat (good fat) can raise the level of high-density lipoproteins (HDL), which carry cholesterol from the heart to the liver and thereby eliminate excess cholesterol; in contrast, consuming saturated fat (bad fat) can raise the level of low-density lipoproteins (LDL), which carry cholesterol from the liver to the rest of the body and thereby lead to blocked coronary arteries. (adapted from Harvard School of Public Health Nutrition Source 2004)

Then, participants were told that each of the three new brands of crackers contained 13 grams of total fat per serving. One brand (the healthy alternative) was portrayed as containing 11 grams of good fat and 2 grams of bad fat, and another (the unhealthy alternative) was portrayed as containing 2 grams of good fat and 11 grams of bad fat. A third, "control" version was portrayed as containing 6.5 grams of both types of fat. All participants were exposed to information about all three versions of the crackers in a counterbalanced manner.

Dependent and other variables. After exposure to information about the fat content of the three cracker versions, participants were asked to rate "How tasty do you think these crackers would be?" and "How much do you think you would enjoy eating them?" on a ten-point scale (1 = "not at all," 10 = "very") for each cracker version. To test whether the influence of the unhealthy = tasty intuition can occur implicitly, we assessed explicitness of participants' beliefs in the intuition by eliciting participants' agreement with the same two statements we used in Experiment 1. Finally, participants provided information on their liking for cheddar-flavored snack crackers and also indicated their perceptions of the unhealthiness of the three cracker versions.

Results

We dropped 4 participants from the analyses for indicating that they did not like any type of cheddar-flavored snack cracker (a rating of 1 on the scale that assessed liking). This left 106 participants (48 men and 58 women).

Manipulation check. We first wanted to assess whether our unhealthiness manipulation worked as intended. We regressed perceived unhealthiness of crackers on their purported bad fat content. The results revealed a positive monotonic relationship between purported bad fat content and perceived unhealthiness ($\beta = +.39$, $t(101) = 9.32$, $p < .001$), revealing that the unhealthiness manipulation worked as we intended.

Taste inferences. The items used to measure tastiness and enjoyment of the crackers were highly correlated (Cronbach's $\alpha = .86$); thus, we averaged them to form a tastiness index. The items used to measure explicitness of belief in the intuition were also highly correlated (Cronbach's $\alpha = .73$); thus, we averaged them to form an explicitness-of-belief index, as in Experiment 1. Using the taste inference index of the crackers as the dependent variable, we conducted a mixed analysis of variance (ANOVA) with three levels of portrayed unhealthiness of the cracker (healthy, control, unhealthy) as a repeated factor and explicitness of belief in the intuition as a continuous between-subjects factor. The results revealed a marginal main effect of unhealthiness ($F(2, 103) = 2.44$, $p = .09$). As we predicted, mean inferred tastiness was highest for the cracker portrayed as containing the highest amount of bad fat ($M = 7.33$) and lowest for the cracker portrayed as containing the lowest amount of bad fat ($M = 5.16$), with the control cracker falling in between ($M = 6.38$). Although the differences between the inferred tastiness of the unhealthy and the control cracker and those between the healthy and the control cracker were not significant ($F(1, 104) = 1.53$ and 1.84, respectively, both $ps > .12$), the focal contrast between tastiness ratings of the unhealthy and healthy crackers was significant ($F(1, 104) = 17.02$, $p < .001$). These results offer evidence consistent with H_1 and indicate that consumers infer the taste of food products when information relevant to assessing their healthiness is provided in a manner consistent with the unhealthy = tasty intuition.

Role of explicitness of belief. The results of the ANOVA also revealed a significant unhealthiness × explicitness-of-belief interaction (F(2, 103) = 12.12, $p < .001$), suggesting that the inference that crackers with more bad fat tasted better was made to differing degrees, depending on participants' strength of belief in the unhealthy = tasty intuition. To test whether the unhealthy = tasty intuition influenced the taste inferences of participants who did not report explicit belief in it, we assessed the mean difference in inferred tastiness ratings of the unhealthy versus healthy cracker at each level of reported explicitness of belief in the intuition (Irwin and McClelland 2001). The results revealed that this difference was significant at every level of belief. That is, as we depict in Figure 2, at all reported levels of disagreement and agreement with the intuition, the difference in inferred tastiness of the unhealthy versus healthy cracker was significantly different from zero (all t's > 2.68, all *ps* < .01).[3]

Discussion

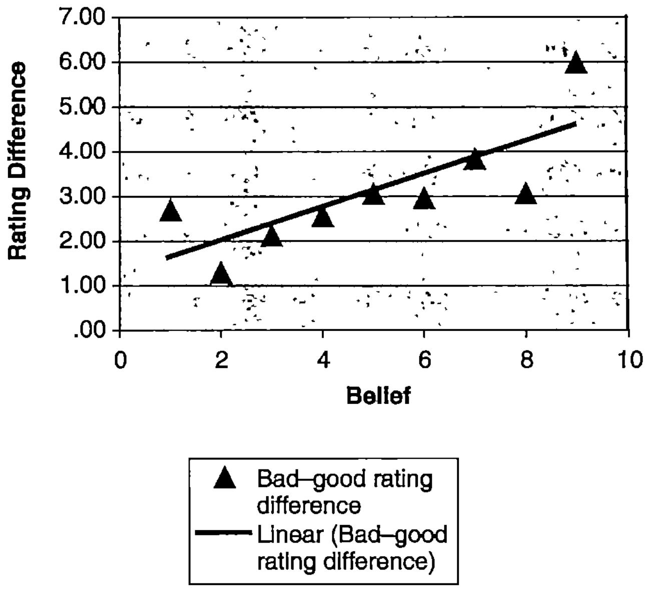
The results from Experiment 2 lead us to conclude that the unhealthy = tasty intuition influences taste inferences. Although participants in the experiment did not taste any of the crackers, and the total amount of fat was the same across the different versions of the cracker, they readily assumed that crackers with more bad fat would taste better. Furthermore, we also found that though the effect of portrayed healthiness on taste inferences was particularly strong among those who reported an explicit belief in the unhealthy = tasty intuition, we also obtained the effect among participants who reported disagreement with the intuition. This suggests that consumers may implicitly subscribe to the view that healthiness and tastiness are inversely correlated, thus conceptually corroborating the results we obtained in Experiment 1.

Experiment 3: Enjoyment of Mango Lassi

We wanted to expand on our findings in several ways in Experiment 3. First, we attempted to address a potential limitation arising from the within-subjects design used in Experiment 2. Because participants were asked to infer the tastiness of three different cracker versions that varied only in terms of portrayed healthiness, it could be argued that the participants may have guessed our hypotheses. Although the significance of the two-way unhealthiness × explicitness-of-belief interaction suggests that demand effects cannot fully explain our findings, we wanted to act conservatively and conceptually replicate these findings using a between-subjects manipulation of unhealthiness. Second, we wanted to demonstrate that the portrayed healthiness of food affects actual (versus inferred) enjoyment (see Figure 1, Stage 2). Such a finding would indicate that the unhealthy = tasty intuition acts as a hypothesis that participants confirm after sampling stimuli (e.g., Hoch and Ha 1986; Shiv, Carmon, and Ariely 2005). Third, we wanted to assess whether the implicit nature of the influence of the unhealthy = tasty intuition can be replicated in the context of actual enjoyment. Such a finding would further attest to the robustness of H_4. Finally, we tested the generalizability of our results by using a different (1) product category, (2) unhealthiness manipulation, and (3) experimental setup.

Forty adults were invited to a housewarming party at the residence of one of the authors of this article. In the invitation to the event (sent by e-mail), invitees were informed that refreshments would be served at the party. The refreshment item targeted for testing our predictions was Mango Lassi (an East Indian delicacy similar to a milkshake). We used the context of the housewarming party, which provided a more realistic setting than the laboratory, to test our predictions.

FIGURE 2
Experiment 2: Difference in Inferred Tastiness of an Unhealthy Versus a Healthy Cracker at Each Level of Explicitness of Belief in the Unhealthy = Tasty Intuition



---

[3] To ensure that our results hold when the unhealthy = tasty intuition is measured as such (i.e., that unhealthy implies tasty rather than that healthy implies not tasty), we conducted an additional study with 287 undergraduate students that reversed the wording of the explicitness-of-belief measure. That is, participants were asked to rate their agreement with the statement "The more unhealthy a food, the better it tastes" rather than the statement "There is no way to make food healthier without sacrificing taste." The results were nearly identical to those we found in Experiment 2. Participants rated the unhealthy cracker (M = 7.31) as significantly tastier than the healthy cracker (M = 5.33; F(1, 286) = 163.71, $p < .0001$), and this difference in taste ratings interacted with explicitness of belief (F(1, 285) = 10.99, $p < .001$), such that higher levels of belief were associated with greater differences in taste ratings. Furthermore, when we used Irwin and McClelland's (2001) method, the difference in rated taste between the unhealthy and the healthy cracker was significantly different from zero at every level of explicit belief in the intuition (all t's > 2.57, all *ps* < .05) except at the lowest level. At the lowest level of explicit belief, this difference approached significance (t = 1.55, $p = .12$).

Participants, Stimuli, and Procedure

At the party, invitees were told that snacks had been ordered from a newly opened Indian restaurant (Hyderabad Masala) and that the operators of this restaurant were interested in catering mainly to an American audience. To lend credibility to the cover story, during appropriate moments at the party, the host nonchalantly expressed that "when the restaurant operators realized that I was going to have a predominantly American crowd at the party, they wondered if the guests would be willing to fill out a short questionnaire about the food. I didn't see any harm in agreeing, especially since they promised to give you a 25% discount on your next visit to the restaurant." A coupon for "25% off the next meal at Hyderabad Masala" was made available to the invitees. To lend further credibility to the cover story, a confederate (a graduate student of East Indian origin) was hired to play the role of an agent representing the restaurant. The confederate handed out the one-page questionnaire that provided information about the following food items offered at the party and elicited participants' judgments of the tastiness and enjoyment of the following items using a ten-point scale (1 = "not at all," 10 = "very"): (1) Idlis, (2) Samosas, and (3) Mango Lassi.[4] These items were placed on a dining table in the formal dining area. The Idlis and Samosas were served on paper plates, and the Mango Lassi was served in prefilled 12-ounce plastic cups.

Because the invitees arrived at various times during a fairly large (four hours) time interval and because the dining area was physically distant from the family room, where people congregated, the invitees tended to be separated from the other guests at the time of sampling the food items and filling out the questionnaire. This minimized potential response bias effects due to social influence (Raghunathan and Corfman 2006). To encourage candidness further, the confederate (who was always present next to the food) stressed to every invitee, "Since we [the restaurant operators] are interested in catering to the American taste, we request that you respond as candidly as possible. Please don't worry about hurting our feelings."

Unhealthiness manipulation. Across all participants, the Idli was portrayed as a healthy item, and the Samosa was portrayed as an unhealthy one. These two items were not the focus of this experiment, but we included them to lend credibility to the cover story and to mask the main target of this experiment (the Mango Lassi). For roughly half of the participants, the Mango Lassi was portrayed as a healthy item ("made out of real mango pulp and milk; generally considered very healthy"), and for the other half, it was portrayed as an unhealthy one ("made out of real mango pulp and milk; generally considered unhealthy"). Note that this manipulation (unlike the one we used in Experiment 2) serves to alter perceptions of unhealthiness directly (rather than manipulate it by altering the content of an ingredient considered unhealthy).

Measurement of explicitness of belief in the intuition. Two days after the party, the host e-mailed the invitees individually, attaching an article that had recently appeared in the local newspaper. The newspaper article suggested that the "low-carb fad" was waning. The purported purpose of the e-mail was to alert the recipients to the content of the "interesting" article, but the true purpose was to gauge the invitees' explicitness of belief in the unhealthy = tasty intuition. Using the news article as a backdrop, the host assessed the invitees' beliefs in the relationship between healthiness and tastiness by asking the following question, "On a scale of one to nine, how much would you agree with the following statement: 'Food that is unhealthy generally tastes better'?" We took invitees' ratings as a measure of the explicitness of their beliefs in the unhealthy = tasty intuition; higher numbers indicated greater explicitness of belief.

Results

About a week after the party, the invitees were individually e-mailed and asked if they had suspected that the true purpose of the questionnaire administered during the party was different from the purported one. We dropped data from one individual who reported suspicion.

Taste inferences. Tastiness and enjoyment of the Mango Lassi were highly correlated with each other (Cronbach's $\alpha = .73$); thus, we averaged them to form an enjoyment index. Using this index as the dependent variable, we conducted a mixed ANOVA with two levels of unhealthiness of Mango Lassi (healthy, unhealthy) as a between-subjects factor and explicitness of belief in the intuition as a continuous between-subjects factor. The results from this analysis appear in Figure 3. Consistent with our prediction (H_2), the mean enjoyment ratings were higher for participants in the condition in which the milkshake was portrayed as being unhealthy ($M = 5.13$) than for participants in the condition in which the milkshake was portrayed as being healthy ($M = 3.30$; $F(1, 35) = 24.29$, $p < .001$).

Role of explicitness of belief. Echoing findings from Experiment 2, the ANOVA also revealed a marginal healthiness × explicitness-of-belief interaction ($F(1, 35) = 3.91$, $p = .06$), indicating that the relationship between enjoyment of the milkshake and its perceived healthiness depended on participants' explicitness of belief in the unhealthy = tasty intuition. Follow-up analyses revealed that the effect of unhealthiness on enjoyment was significant both at one standard deviation above the mean belief level ($t(1, 35) = 4.84$, $p < .001$) and at one standard deviation below it ($t(1, 35) = 2.00$, $p = .05$), indicating that portrayed healthiness of food significantly influenced participants' enjoyment of the food, even among those who did not report agreeing with the unhealthy = tasty intuition. These results provide further support for our hypothesis that the influence of the intuition on consumption decisions can take place outside of awareness.

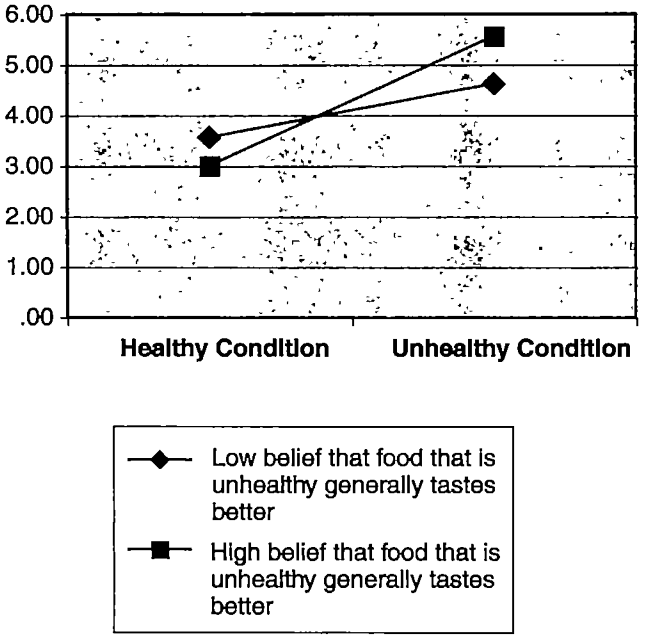
Discussion

Viewed in conjunction with the results from Experiments 1 and 2, the results from Experiment 3 indicate that taste

---

[4] Idli is a traditional South Indian dish made by steaming a mixture of mainly lentils and rice. Samosas are deep-fried vegetable dumplings. Lassi refers to sweet drinks made with a base of yogurt.

FIGURE 3
Experiment 3: Mango Lassi Enjoyment Ratings × Explicitness of Belief That Food That Is Unhealthy Generally Tastes Better



inferences made on the basis of perceived unhealthiness are confirmed on sampling the food, implying that a mechanism akin to the hypothesis confirmation bias may underlie the prevalence of the unhealthy = tasty intuition. Furthermore, the finding that even participants who reported disagreeing that healthiness needs to be sacrificed in order to enhance tastiness enjoyed the milkshake more when it was portrayed as unhealthy (versus healthy) indicates that the influence of the unhealthy = tasty intuition can be implicit, thus conceptually replicating results from Experiments 1 and 2. Finally, the results were consistent with our predictions even though we used a different (1) design (between- versus within-subjects), (2) product category, and (3) unhealthiness manipulation; this attests to the robustness of the influence of the unhealthy = tasty intuition.

Experiment 4: Choice of Crackers

The results from Experiments 2 and 3 suggest that people's belief in the unhealthy = tasty intuition is sufficiently strong to influence taste inferences and enjoyment of food. The objective in Experiment 4 was to further explore whether the intuition would influence choice decisions (see both Figure 1, Stage 3, and H_3). We know from previous research that when consumers are faced with the task of choosing among food products, better-tasting options have a higher likelihood of being chosen when a hedonic goal is relatively more salient (e.g., Dhar and Simonson 1999; Shiv and Fedorikhin 1999). Extrapolating from these results, we expect that participants will prefer an option portrayed as being more (versus less) unhealthy when a hedonic goal is relatively more salient because according to our hypothesis, people will infer this option to taste better (and thus will judge it as a better candidate for fulfilling the hedonic goal).

Participants, Stimuli, and Procedure

Two hundred ninety-three undergraduate students (107 men, 184 women, 2 unknown gender) participated in this experiment for course extra credit. They were first told that a national manufacturer of crackers was interested in assessing which of two versions of a cheddar-flavored snack cracker is preferred in choice tasks. Subsequently, information about good and bad fat (identical to the information we used in Experiment 1) was provided before participants chose between a cracker version with 11 grams of bad fat and 2 grams of good fat (unhealthy version) and one with 2 grams of bad fat and 11 grams of good fat (healthy version).

Before indicating their choice, approximately half of the participants were told to imagine that they were "craving something really tasty" and that "they wanted to reward themselves with a nice snack." The purpose of these statements was to prime a hedonic, or enjoyment, goal. The remaining participants were not exposed to such a prime; they were simply told to imagine that they were "in the mood for a snack." The dependent variable was which cracker version participants selected in a forced-choice task. We elicited participants' explicitness of belief in the unhealthy = tasty intuition using the same questions we used in Experiments 1 and 2. Consistent with H_3, we expected that the relative preference for the unhealthy (versus healthy) cracker version would be higher among participants who were primed with the hedonic goal.

Results and Discussion

Manipulation check and choice. As we expected, participants perceived the cracker with the higher content of good (versus bad) fat content as healthier (Ms = 6.13 and 3.05; t(291) = 1067.71, $p < .01$). To test our prediction that the relative preference for the unhealthy (versus healthy) option would be higher among those primed with a hedonic goal, we conducted a chi-square test. As we predicted, the proportion of participants who chose the unhealthy option was significantly greater among those who were primed with the hedonic goal (M = 21%) than among those who were not primed with this goal (M = 8%; $\chi^2 = 15.88$, $p < .001$; $\phi = .23$).[5]

Role of explicitness of belief. To test the role of explicitness of belief, we conducted an ANOVA using the prime and explicitness of belief in the intuition as the predictor variables and choice as the dependent variable. The results revealed a nonsignificant prime × explicitness-of-belief interaction (F(1, 258) = .59, $p > .44$), indicating that, regardless of whether the participants reported agreeing with the unhealthy = tasty intuition, they made choice decisions as if they did.

---

[5] The proportion of participants who chose the unhealthy cracker was low even in the hedonic prime condition, presumably because participants did not want to indicate that they were interested in consuming a cracker that was explicitly portrayed as being unhealthy.

Thus, the results from Experiment 3 conceptually replicate those we obtained previously, providing evidence that the unhealthy = tasty intuition affects not just inferences and judgments but actual choice decisions as well. We found that when a hedonic goal was made temporarily more salient, participants were more likely to select the unhealthy cracker, presumably because they inferred that it would taste better. Furthermore, the lack of significance of the prime × explicitness-of-belief interaction suggests that the influence of the intuition can take place outside the awareness of consumers, indicating that controlling for the influence of the intuition may be difficult.

General Discussion

Understanding the effects of food consumption habits on health has emerged as a topic of significance in recent years. The question of why consumers overconsume food despite the negative consequences of doing so has especially piqued the interest of academics (e.g., Kozup, Creyer, and Burton 2003; Oakes 2005; Oakes and Slotterback 2005; Seiders and Petty 2004; Smith 2004; Tepper and Trail 1998; Urala and Lähteenmäki 2004; Verbeke 2005; Wansink 2004 a, b; Wansink and Chandon 2006; Wansink and Huckabee 2005). Previous findings are consistent with the idea that people overconsume food that is considered unhealthy either because they are unaware of the negative consequences of doing so (presumably because of poor marketing and/or presentation of nutritional information) or because they lack the willpower to resist such foods. In what we believe is the first rigorous empirical demonstration of its kind, our findings indicate that in addition to these factors, a more subtle force may be at work; namely, consumers overconsume foods that they perceive as unhealthy because they assume that such food tastes better.

The results from Experiment 1, in which we used the well-established IAT procedure (Greenwald, McGhee, and Schwartz 1988), demonstrated not only that people subscribe to the unhealthy = tasty intuition but also that this intuition operates at an implicit level. Both participants who reported an explicit belief in the intuition and those who did not took longer to categorize stimuli in the incongruent (versus congruent) condition, which suggests that the implicit subscription to the unhealthy = tasty intuition is robust.

The results from three other experiments confirmed our predictions that when information pertaining to the assessment of the healthiness of food items is provided, foods perceived as less healthy are (1) inferred to taste better, (2) enjoyed more during actual consumption, and (3) preferred in a choice task when a hedonic goal is more (versus less) salient. We obtained these patterns of results even among those who disagreed that healthiness is inversely related to tastiness, which suggests that the influence of the unhealthy = tasty intuition can take place outside of awareness; this result conceptually replicates Experiment 1's results, which demonstrate that subscription to the unhealthy = tasty intuition is implicit.

Theoretical Implications

Validity of the intuition and reasons for its propagation. In light of our findings, an important question that emerges is whether the unhealthy = tasty intuition is valid, that is, whether the healthiness and tastiness of food items are indeed negatively correlated with each other in the real world. Although the intuition appears to enjoy widespread subscription (e.g., Keller, Sternthal, and Tybout 2002), there is little scientific evidence to support the view that tastiness and healthiness are negatively correlated with each other. Indeed, from an evolutionary standpoint, evidence points to the opposite, namely, that tastiness and healthiness are positively correlated with each other. For example, in general, it is accepted among researchers (e.g., Drewnowski 1997; Smith 2004) that certain food groups (e.g., carbohydrates, fat) are perceived as tastier precisely because these foods have proved effective for survival; in these researchers' view, tastiness has served humans well as a proxy for healthiness.

Why then do negative health consequences result from consuming the food groups (e.g., carbohydrates, fat) that are, evolutionarily speaking, considered healthy? Medical evidence suggests that it is the overconsumption, not moderate levels of consumption, of these food groups that leads to negative health consequences (e.g., Drewnowski 1997).[6] Indeed, consuming fat and carbohydrates in moderation has been found to have salutary effects on health (Better Health Channel 2004). Consistent with this view, findings indicate that overconsumption of any food group can result in negative health consequences.

There is yet another reason to question the validity of the unhealthy = tasty intuition. Note that to assess the actual relationship between healthiness and tastiness, it is first necessary to measure the healthiness and tastiness of various food options accurately. We believe that this is a deceptively difficult task because neither judgments of healthiness nor those of tastiness are likely to be stable across time or across people. Oakes's (2005) findings demonstrate how judgments of healthiness are notoriously susceptible to contextual influences. In one study, for example, Oakes finds that a lower-calorie snack portrayed as being "disreputable" in terms of health (i.e., a Snickers miniature bar with 47 calories) was inferred to promote greater weight gain than one portrayed as being more "reputable" in terms of health (i.e., a half cup of green peas with 62 calories). Thus, participants in this study relied on their subjective perceptions of the snacks' healthiness rather than on more diagnostic information (the snacks' caloric value) to infer their potential for weight gain. There is reason to believe that judgments of tastiness may be similarly susceptible to contextual influences. For example, Carpenter and Nakamoto's (1989) findings suggest that what people happen to taste first (e.g., Coke or Pepsi) may become their preferred option. An overview of other relevant findings suggests that there is little agreement (among adults) about tastiness ratings of food (see, e.g., Drewnowski 1997, p. 240).

---

[6] Needless to say, underconsumption of nutrients can also lead to negative health consequences. Indeed, malnourishment is the world's deadliest cause of premature death (World Health Organization 2002). However, because this research is restricted to the context of the United States, where food is plentiful, we restrict our discussion to negative consequences of overconsumption.

Nonconscious influence. It is well known that consumers' decisions are influenced by their lay theories and intuitions even when the lay theories are invalid (Broniarczyk and Alba 1994; Shiv, Carmon, and Ariely 2005). What is noteworthy in the current research is that our predictions were confirmed across two different types of food products, three types of judgment tasks, and two conceptually different types of unhealthiness manipulations. The generality of our findings across these experimental differences hints at the prevalence and strength with which people believe in the unhealthy = tasty intuition. Even more intriguing, however, is the indication of the implicit nature of the influence of the intuition. The results from Experiment 1 confirmed that people may have implicitly learned to correlate unhealthiness with tastiness. The results from Experiments 2 and 3 offer conceptual replication by showing that even participants who reported disagreeing with the view that healthiness and tastiness are inversely correlated behaved as if they did. These findings indicate that the influence of the intuition may lie outside the awareness of consumers making judgments about what to consume. At a conceptual level, this notion shares similarities with Lewicki and colleagues' (e.g., Lewicki 1986; Lewicki and Hill 1989; Lewicki, Hill, and Czyzewska 1997) work on nonconscious learning. The idea that a person's judgments and decisions may be influenced by factors outside his or her awareness is also echoed in various other findings, such as mere exposure effects (Zajonc 1980), truth effects (Hawkins and Hoch 1992), and assimilation/contrast effects (e.g., Schwarz and Bless 1992).

Managerial, Consumer, and Public Policy Implications

Although the obesity problem afflicting people in the United States may be due, in part, to the ubiquity of food high in fat content, our findings indicate that even if options with less fat were available, consumers may choose to consume foods high in fat because of a misguided notion that they need to eat unhealthy to eat tasty or that eating healthy means having food that is not tasty. Evidence from food scientists suggests that despite current concerns about healthy eating, health is secondary to taste in selecting many food products (for a discussion of taste versus health in consumer selection of corn chips, see Tepper and Trail 1998). There is also evidence in recent research on functional food—a new category of products that promise consumers improvements in targeted physiological functions, such as lowering cholesterol (Urala and Lähteenmäki 2004)—that consumers are unwilling to compromise on taste even for the sake of the health benefits these foods offer (Verbeke 2006). This lesson was perhaps most starkly revealed in the interesting story of the McLean, a low-fat burger that McDonald's tried to market in the early 1990s. Although the burger was actually preferred to burgers that contained more fat in blind taste tests, it was a miserable market failure (Gladwell 2001).

Given our evidence that the influence of the unhealthy = tasty intuition is strong and may take place outside the consumer's awareness, this intuition could be a major contributing factor to obesity in consumers. Therefore, a key question is, What (if anything) can be done to counteract its potentially devastating impact on consumer food choice decisions and, subsequently, on consumer health?

More research is being directed at understanding both how to market products such that consumers make better food choices (Andrews, Netemeyer, and Burton 1998; Kozup, Creyer, and Burton 2003; Seiders and Petty 2004; Wansink and Huckabee 2005) and how to help consumers make changes in their environment to help control their food consumption (Wansink 2004a). These streams of research offer guidance to the food industry, which has been accused of contributing to the obesity epidemic, and to consumers who want to make healthier choices that are still enjoyable, despite the pervasiveness of a belief in the unhealthy = tasty intuition. Possibilities for controlling the effect of the unhealthy = tasty intuition (and its potential for causing negative health consequences) include (1) controlling the volume of unhealthy but tasty food that is eaten, (2) changing unhealthy foods to make them less unhealthy but still tasty, and (3) providing consumers with better information about what constitutes "healthy." We address each of these possibilities in the remainder of this discussion.

Limiting consumption volume of unhealthy foods. If consumers infer that unhealthy foods are tastier, they will possess a strong desire to eat foods that are relatively unhealthy (e.g., foods that are high in fat or high in calories). A possible solution to this problem would be for both marketers and consumers to take steps to ensure that if these foods are going to be eaten, they are eaten in reasonable quantities. Such steps are important because, as we mentioned previously, medical evidence suggests that it is the overconsumption, not moderate levels of consumption, of foods considered unhealthy that leads to negative health consequences (e.g., Drewnowski 1997). Therefore, measures aimed at restricting consumption of foods considered unhealthy (but tasty) to moderate levels can go a long way toward counteracting the potential negative health consequences of a belief in the unhealthy = tasty intuition.

Responsible marketers and individual consumers can take such steps to control consumption volume of unhealthy foods. For example, Wansink and Huckabee (2005) suggest that marketers can guide consumers' volume decisions by changing packaging (e.g., offering multipacks with smaller individual servings, offering premium-priced smaller packaging), changing the volume of implied serving sizes in packaging information (e.g., designating a serving as three cookies versus six), and providing consumers with easier ways to assess how much they are consuming (e.g., placing markers on the outsides of packaging to indicate a serving size, placing indicator lines inside packaging). Consumers can also make changes to their personal environment to help reduce consumption.

Wansink (2004a) offers many practical alterations that can be made to both the eating and the food environments, including pre-serving portions when snacking while distracted (e.g., while watching television); repackaging foods into small containers to suggest smaller consumption norms; never eating directly from a package; and using smaller plates, bowls, and eating utensils. Such steps can allow marketers to continue to market unhealthy foods in responsible ways and allow consumers to continue to enjoy the foods they consider tasty, albeit in moderation. Although many of these changes seem small, Wansink and Huckabee (2005) point out that 80% of the population gain weight because of a calorie excess of less than 50 calories a day. The changes suggested here could allow consumers to enjoy the foods they consider tasty in smaller, controlled quantities, perhaps enough to avoid consumption of those additional 50 calories.

Reformulating unhealthy foods. Another possibility for controlling the effects of the unhealthy = tasty intuition is to change the composition of unhealthy foods. One alternative is to reformulate high-energy-density foods (i.e., foods high in calories) to lower their energy density (i.e., calorie content) by replacing some of the fat with water, fiber filler, or vegetables (Wansink and Huckabee 2005). Although evidence from our studies suggests that people believe that food cannot be made healthier without sacrificing taste (and, thus, that these lower-energy-density foods would be less appetizing), taste tests indicate otherwise. Rolls, Ello-Martin, and Tohill's (2004) research suggests that consumers who eat lower-energy-density foods eat the same volume as they would of the unmodified food (resulting in fewer calories consumed), rate themselves as equally satisfied as those who eat full-calorie dense foods, and, most important in our research context, do not perceive the foods as tasting worse (Kral and Rolls 2004; Rolls 2005; Wansink and Huckabee 2005). Research suggests that up to 20% of the fat in a high-energy-density food can be replaced with low-density items (e.g., fruits, vegetables) without consumers noticing a difference in taste (Rolls, Ello-Martin, and Tohill 2004).

Such findings suggest that many consumers are not knowledgeable enough about food composition to realize that the unhealthy foods they love can be successfully reformulated to be just as tasty without being as unhealthy. Because such modifications could have a large impact on the number of calories consumed, they could be very beneficial if consumers are educated to believe that the healthier version of their favorite unhealthy food is just as tasty. However, because of the pervasiveness of the unhealthy = tasty intuition, such a strategy could backfire (i.e., despite research evidence suggesting the taste had not truly changed, consumers could respond to such a modification as they did in Experiment 3, in which they allowed the knowledge that a food is now healthier to color their actual taste perceptions).

To combat this problem, an obvious solution would be to stress both taste and health for these items, with the taste dimension being the dominant dimension. By stressing the taste dimension first, companies could work on breaking down the unhealthy = tasty intuition and increasing the probability of choice. Marketing strategies involving sampling, credible sources, and opinion leaders would be critical. First, sampling would be a key factor in gaining acceptance for these new, healthier products because consumers will not believe that they are tasty unless they have tried the product themselves. A free sample represents a low-risk way to experience the taste. These samples could be distributed through the mail or, more appropriately, handed out in health clubs or at sporting/fitness events.

Second, well-respected health or fitness experts could be employed to endorse these products. To change engrained beliefs, a credible source is needed. Consumers are more likely to accept a message about a product when the source is considered trustworthy and viewed as an expert on the topic (e.g., Goldsmith, Lafferty, and Newell 2000; Sternthal, Dholakia, and Leavitt 1978). This technique could be particularly effective if these sources are opinion leaders. For example, health club employees and trainers are often looked to by consumers for tips on exercise and diet. A marketing or public policy program could educate and persuade these people on the tastiness and health benefits of the new products. They could then be important transmitters of this information to consumers. Thus, this message would be more likely to be accepted and acted on.

Another possibility might be to stimulate "buzz" for certain products among consumers. If marketers could convince a key group of consumers in the word-of-mouth network that their food products taste good and are part of a healthy diet, this group could influence and excite other groups of consumers. A similar phenomenon occurred in the fast and widespread acceptance of the Atkins diet, in which consumers were willing to give up foods high in carbohydrates for quick weight loss.

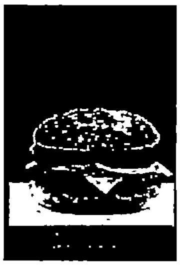
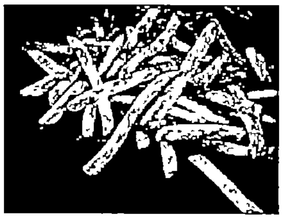
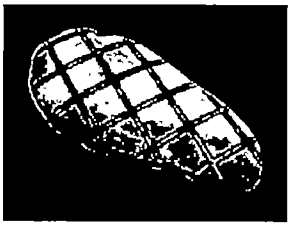
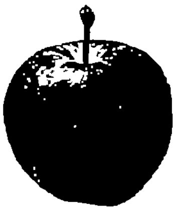
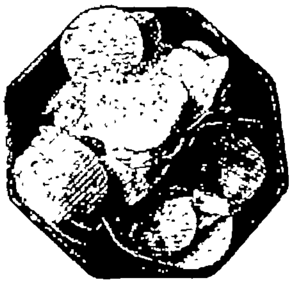
Reeducating consumers about what is "healthy." A final avenue for counteracting the effects of the unhealthy = tasty intuition is to educate consumers about what constitutes healthy. Determining what is healthy or unhealthy is a deceptively difficult task because, as we discussed previously, healthiness is not an inherent quality of food; rather, it depends on the conditions (e.g., quantity, genetic makeup of the person consuming it) under which it is consumed. Nevertheless, consumers can be educated to approach nutrition in a more balanced fashion than they currently do. In Oakes's (2005) study, which we described previously, consumers overemphasized fat content when they assessed the healthiness of foods and relied on a categorization system (of foods as bad versus good) using stereotypes about types of food/ingredients to determine how much weight a food would cause someone to gain. Had they relied instead on a more complete nutritional assessment of the foods, they would have found that the food they considered healthy on the basis of fat content (peas) would actually lead to more weight gain than the food they considered unhealthy (Snickers) because of the caloric content of the two foods.

What can be done to convince consumers to consider all types of nutritional information? A possibility is to develop an easy-to-use nutrition index that combines the various types of relevant information into a single number, which consumers could then use to make easy comparisons across different types of foods. This would enable consumers to choose healthier food with minimal effort and confusion. Furthermore, reeducating consumers to evaluate the healthiness of a food could actually change consumers' perceptions about what constitutes healthy and unhealthy.

Limitations and Further Research

An interesting implication of our main hypothesis (that people subscribe to the unhealthy = tasty intuition) is that regardless of which food group is associated with unhealthiness, food items that are higher (versus lower) in that ingredient will be inferred to be tastier. Thus, for example, in contexts in which carbohydrates are considered unhealthy (as is likely to be the case among those following the Atkins diet), food items higher (versus lower) in carbohydrates will be (1) inferred to taste better, (2) enjoyed more, and (3) preferred when a hedonic goal is more (versus less) salient. Indeed, the results from Experiment 3, in which no mention was made of the particular ingredients

APPENDIX
Experiment 1: Stimuli Used in IAT

		Word List	
Unhealthy Foods	**Healthy Foods**	**Enjoyable**	**Not Enjoyable**
		Tasty Delicious Yummy Appetizing Flavorful Appealing Mouthwatering	Disliked Less tasty Unappealing Bland Flavorless Unappetizing Unpalatable
			
			
			

that made the Mango milkshake more or less healthy, offer conceptual support for this hypothesis.

An important related empirical question that is still unanswered by the current research is whether other manipulations of unhealthiness would produce similar results in terms of taste inferences, enjoyment, and choice. For example, would altering the content of preservatives (assuming that they are considered unhealthy) alter perceptions of tastiness in a manner consistent with the unhealthy = tasty intuition? If not, identifying the manipulations that lead to the types of results we document in this research would be important for both theoretical and substantive reasons. Similarly, further research should ascertain the boundaries of the content of unhealthy ingredients within which results consistent with the unhealthy = tasty intuition will be obtained. Is there a limit beyond which increasing unhealthiness does not lead to enhanced inferences of tastiness? Levin and Gaeth's (1988) results suggest that there is. They find that beef that was portrayed as 25% fat was expected to taste worse than beef that was portrayed as 75% lean. These results indicate that there is a range within which increasing the portrayed unhealthiness of food leads to enhanced perceptions of tastiness but also a range beyond which it does not.

A final and interesting extension of our work would be to assess whether and to what extent socioeconomic and cultural factors moderate our effects. In contexts in which food is not plentiful (e.g., in developing countries, such as India or China, and in underdeveloped countries, such as Somalia or Cambodia), people may believe tastiness is positively correlated with healthiness, as evolution intended (Drewnowski 1997; Smith 2004). If so, we may expect to find a pattern of results that is opposite to that which we find in this research. Similarly, in epicurean cultures (e.g., France), in which food is valued primarily for its tastiness, it is possible that the perception of an inverse relationship between healthiness and tastiness is held with less conviction.

REFERENCES

Allen, Chris T. and Chris A. Janiszewski (1989), "Assessing the Role of Contingency Awareness in Attitudinal Conditioning with Implications for Advertising Research," *Journal of Marketing Research*, 26 (February), 30–43.

Andrews, J. Craig, Richard G. Netemeyer, and Scot Burton (1998), "Consumer Generalization of Nutrient Content Claims in Advertising," *Journal of Marketing*, 62 (October), 62–75.

Berry, Christopher J. (1994), *The Idea of Luxury*. Cambridge, UK: Cambridge University Press.

Better Health Channel (2004), "Sugar," (accessed February 12, 2004), [available at http://www.betterhealth.vic.gov.au/bhcv2/bhcarticles.nsf/pages/Sugar?OpenDocument].

Bettman, James R., Mary Frances Luce, and John W. Payne (1998), "Constructive Consumer Choice Processes," *Journal of Consumer Research*, 25 (December), 187–217.

Broniarczyk, Susan M. and Joseph W. Alba (1994), "The Role of Consumers' Intuitions in Inference Making," *Journal of Consumer Research*, 21 (December), 393–407.

Brownell, Kelly D. and Katherine Battle Horgen (2003), *Food Fight: The Inside Story of the Food Industry, America's Obesity Crisis, and What We Can Do About It*. New York: McGraw-Hill.

Carpenter, Gregory S. and Kent Nakamoto (1989), "Consumer Preference Formation and Pioneering Advantage," *Journal of Marketing Research*, 26 (August), 285–98.

Deighton, John (1984), "The Interaction of Advertising and Evidence," *Journal of Consumer Research*, 11 (December), 763–70.

Dhar, Ravi and Itamar Simonson (1999), "Making Complementary Choices in Consumption Episodes: Highlighting Versus Balancing," *Journal of Marketing Research*, 36 (February), 29–44.

Drewnowski, Adam (1997), "Taste Preferences and Food Intake," *Annual Review of Nutrition*, 17, 237–53.

Food Marketing Institute (1996), *Trends in the United States: Consumer Attitudes and the Supermarket*. Chicago: Food Marketing Institute.

Gladwell, Malcolm (2001), "The Trouble with Fries: Fast Food Is Killing Us. Can It Be Fixed?" *The New Yorker*, (March 5), (accessed February 12, 2004), [available at http://www.gladwell.com/2001/2001_03_05_a_fries.htm].

Goldsmith, Ronald E., Barbara A. Lafferty, and Stephan J. Newell (2000), "The Impact of Corporate Credibility on Consumer Reactions to Advertisements and Brands," *Journal of Advertising*, 29 (3), 43–54.

Greenwald, Anthony G., Debbie E. McGhee, and Jordan L.K. Schwartz (1998), "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test," *Journal of Personality and Social Psychology*, 74 (6), 1464–80.

———, Brian A. Nosek, and Mahzarin R. Banaji (2003), "Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm," *Journal of Personality and Social Psychology*, 85 (2), 197–216.

Harvard School of Public Health Nutrition Source (2004), "Fats and Cholesterol: The Good, the Bad, and the Healthy Diet," (accessed February 12, 2004), [available at http://www.hsph.harvard.edu/nutritionsource/fats.html].

Hawkins, Scott A. and Stephen J. Hoch (1992), "Low Involvement Learning: Memory Without Evaluation," *Journal of Consumer Research*, 19 (September), 212–25.

Hoch, Stephen J. and Young-won Ha (1986), "Consumer Learning: Advertising and the Ambiguity of Product Experience," *Journal of Consumer Research*, 13 (September), 221–33.

Irwin, Julie R. and Gary H. McClelland (2001), "Misleading Heuristics and Moderated Multiple Regression Models," *Journal of Marketing Research*, 38 (February), 100–109.

Keller, Kevin Lane, Brian Sternthal, and Alice Tybout (2002), "Three Questions You Need to Ask About Your Brand," *Harvard Business Review*, 80 (September), 80–85.

Kelly, Harold H. and John W. Thibaut (1969), "Group Problem Solving," in *The Handbook of Social Psychology*, 2d ed., Gardner Lindsey and Elliot Aronson, eds. Reading, MA: Addison-Wesley, 1–101.

Kivetz, Ran and Itamar Simonson (2002), "Earning the Right to Indulge: Effort as a Determinant of Customer Preferences Toward Frequency Program Rewards," *Journal of Marketing Research*, 39 (May), 155–70.

Kozup, John C., Elizabeth H. Creyer, and Scot Burton (2003), "Making Healthful Food Choices: The Influence of Health Claims and Nutrition Information on Consumers' Evaluations of Packaged Food Products and Restaurant Menu Items," *Journal of Marketing*, 67 (April), 19–34.

Kral, Tanja V. and Barbara J. Rolls (2004), "Energy Density and Portion Size: Their Independent and Combined Effects on Energy Intake," *Physiology and Behavior*, 82 (1), 131–38.

Levin, Irwin P. and Gary J. Gaeth (1988), "How Consumers Are Affected by the Framing of Attribute Information Before and After Consuming the Product," *Journal of Consumer Research*, 15 (December), 374–78.

Lewicki, Pawel (1986), "Processing Information about Covariations That Cannot Be Articulated," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12 (January), 135–46.

——— and Thomas Hill (1989), "On the Status of Nonconscious Processes in Human Cognition: Comment on Reber," *Journal of Experimental Psychology. General*, 118 (3), 239–41.

———, ———, and Maria Czyzewska (1997), "Hidden Covariation Detection: A Fundamental and Ubiquitous Phenomenon," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23 (January), 221–28

Maslow, Abraham (1962), *Toward a Psychology of Being*. Oxford: Van Nostrand.

Morris, Michael W., Tanya Menon, and Daniel R. Ames (2001), "Culturally Conferred Conception of Agency: A Key to Social Perception of Persona, Groups, and Other Actors," *Journal of Personality and Social Psychology*, 5 (April), 169–82.

Nestle, Marion (2003), *Food Politics: How the Food Industry Influences Nutrition and Health*. Berkeley: University of California Press.

Oakes, Michael E. (2005), "Beauty or Beast: Does Stereotypical Thinking About Foods Contribute to Overeating?" *Food Quality and Preference*, 16 (July), 447–54.

——— and Carole S. Slotterback (2005), "Too Good to Be True: Dose Insensitivity and Stereotypical Thinking of Foods' Capacity to Promote Weight Gain," *Food Quality and Preference*, 16 (December), 675–81.

Parloff, Roger (2003), "Is Fat the Next Tobacco?" *Fortune*, (September), 50–54.

Raghunathan, Rajagopal and Kim Corfman (2006), "Is Happiness Shared Doubled and Sadness Shared Halved? Social Influence on Enjoyment of Hedonic Experiences," *Journal of Marketing Research*, 43 (August), 386–94.

Rolls, Barbara (2005), *The Volumetrics Eating Plan: Techniques and Recipes for Feeling Full on Fewer Calories*. New York: HarperCollins.

———, Julie A. Ello-Martin, and Beth C. Tohill (2004), "What Can Intervention Studies Tell Us About the Relationship Between Fruit and Vegetable Consumption and Weight Management?" *Nutrition Reviews*, 62 (1), 1–17.

Ross, Lee and Richard E. Nisbett (1991), *The Person and the Situation: Perspectives of Social Psychology*. New York: McGraw-Hill.

Schlosser, Eric (2001), *Fast Food Nation: The Dark Side of the All-American Meal*. Boston: Houghton Mifflin.

Schwarz, Norbert and Herbert Bless (1992), "Constructing Reality and Its Alternatives: An Inclusion/Exclusion Model of Assimilation and Contrast Effects in Social Judgment," in *The Construction of Social Judgments*, Leonard L. Martin and Abraham Tesser, eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 217–45.

Seiders, Kathleen and Ross D. Petty (2004), "Obesity and the Role of Food Marketing. A Policy Analysis of Issues and Remedies," *Journal of Public Policy & Marketing*, 23 (Fall), 153–69.

Shiv, Baba, Ziv Carmon, and Dan Ariely (2005), "Placebo Effects of Marketing Actions: Consumers May Get What They Pay For," *Journal of Marketing Research*, 42 (November), 383–93.

——— and Alexander Fedorikhin (1999), "Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making," *Journal of Consumer Research*, 26 (December), 278–92.

Smith, Trenton G. (2004), "The McDonald Equilibrium: Advertising, Empty Calories, and the Endogenous Determination of Dietary Preferences," *Social Choice and Welfare*, 23 (3), 383–413.

Sternthal, Brian, Ruby R. Dholakia, and Clark Leavitt (1978), "The Persuasive Effect of Source Credibility: A Situation Analysis," *Public Opinion Quarterly*, 42 (Fall), 285–314.

Sujan, Mita and Christine Dekleva (1987), "Product Categorization and Inference Making: Some Implications for Comparative Advertising," *Journal of Consumer Research*, 14 (December), 372–78.

Tepper, Beverly J. and Amy C. Trail (1998), "Taste or Healthy: A Study on Consumer Acceptance of Corn Chips," *Food Quality and Preference*, 9 (4), 267–72.

Urala, Nina and Lisa Lähteenmäki (2004), "Attitudes Behind Consumers' Willingness to Use Functional Foods," *Food Quality and Preference*, 15 (October–December), 793–803.

Verbeke, Wim (2006), "Functional Foods: Consumer Willingness to Compromise on Taste for Health?" *Food Quality and Preference*, 17 (1–2), 126–31.

Wansink, Brian (2004a), "Environmental Factors That Increase the Food Intake and Consumption Volume of Unknowing Consumers," *Annual Review of Nutrition*, 24, 455–79.

——— (2004b), *Marketing Nutrition*. Champaign: University of Illinois Press.

——— and Pierre Chandon (2006), "Can 'Low Fat' Nutrition Labels Lead to Obesity?" *Journal of Marketing Research*, 43 (November), forthcoming.

——— and Mike Huckabee (2005), "De-Marketing Obesity," *California Management Review*, 47 (Summer), 6–18.

Weber, Max (1998), *The Protestant Work Ethic and the Spirit of Capitalism*, 2d ed. Los Angeles: Roxbury.

World Health Organization (2002), *The World Health Report*, (accessed January 21, 2005), [available at http://www.who.int/whr/2002/en/].

Zajonc, Robert B. (1980), "Feeling and Thinking: Preferences Need No Inferences," *American Psychologist*, 35 (February), 151–75.

Zhang, Jane (2005), "House Passes 'Cheeseburger' Responsibility Bill," *The Wall Street Journal*, (October 20), A4.

Daniel J. Howard & Roger A. Kerin

Broadening the Scope of Reference Price Advertising Research: A Field Study of Consumer Shopping Involvement

The current understanding of reference pricing requires broadening to reflect contemporary advertising practices and shopping behavior. A content analysis of 13,594 newspaper retail advertisements reveals that it is uncommon for reference pricing to be used alone in advertisements: 87.2% of the advertisements featuring reference pricing also include limited-time availability (i.e., "Three days only!") or sale announcements in the headline or copy. Using dual-process theories of persuasion as a conceptual framework, the authors conduct three field experiments that compare the conditions under which the use of limited-time availability and sale announcements in advertisements that feature reference prices affects consumer price perceptions and store shopping intentions. The findings yield three insights for reference price advertising. First, reference pricing results in favorable price perceptions and store shopping intentions only when consumers are shopping for a product. Second, when consumers are shopping for a product, reference pricing paired with limited-time availability results in more favorable price perceptions and store shopping intentions than the use of either technique alone. Third, the use of sale announcements and limited-time availability in reference price advertisements has a favorable effect on price perceptions and store shopping intentions when consumers are not shopping for a product.

Regularly: $599; Now: $359
Sale! Three Days Only

Most consumers have seen variations of one or both of these two headlines in a retail advertisement. The top headline features reference pricing, in which a current, lower price is compared with a former, higher price. The bottom headline features scarcity—specifically, the limited-time availability of the sale offer. The particular wording and juxtaposition of each headline represent information and value cues that are intended to influence consumer price perceptions favorably. Does this information achieve its intended effect, and is there a difference in the effect when consumers are or are not "in the market" for a product? Does the use of redundant value cues improve or have no effect on price perceptions, and what conditions influence this? In this article, we explore these questions and examine why and when commonly used limited-time availability and sale announcements affect consumer attitudes toward product prices and intentions of shopping at stores featured in retail advertisements that use reference prices. In doing so, we attempt to broaden the scope of research on reference price advertising and understand its implications for practice.

Although reference pricing, limited-time availability, and sale announcements appear with some regularity in retail advertising, little empirical evidence documents the incidence and coincidence of their use. For example, a content analysis of sales advertisements by six major retail chains over the 2001 Labor Day weekend revealed that approximately 72% employed reference pricing in their advertisements (Chandrashekaran, Monroe, and Viswanathan 2004). Although not reported, it is reasonable to speculate that many, if not most, of these advertisements also contained headlines or copy heralding a "Three days only!" Labor Day weekend sale to convey the limited-time availability of price reductions offered.

To obtain a broader perspective on the incidence and coincidence of reference pricing, limited-time availability, and sale announcements in retail advertising, we conducted a content analysis of 13,594 retail advertisements in four major, geographically dispersed U.S. newspapers over a six-month period in 2004.[1] We examined the *Chicago Tri-*

Daniel J. Howard is Professor of Marketing (e-mail: dhoward@mail.cox.smu.edu), and Roger A. Kerin is Harold C. Simmons Professor of Marketing (e-mail: rkerin@mail.cox.smu.edu), Edwin L. Cox School of Business, Southern Methodist University. The authors thank Tiffani Oltmanns and Daniel Taraba for their assistance with this project and William Bearden, Richard Briesch, Rajesh Chandrashekaran, Robert Cialdini, and the two anonymous *JM* reviewers for constructive comments on a previous draft of this article.

To read or contribute to reader and author dialogue on this article, visit http://www.marketingpower.com/jmblog.

[1] Sources for all four newspapers confirmed that advertisements that are 17 square inches or larger in size (e.g., two columns wide × four inches in height) account for more than 90% of the display ad revenues, and we examined all such advertisements. We collected data from four newspapers to provide representation of different regions of the country for a 195-day period that included Memorial Day through Thanksgiving. Two research assistants performed the coding, and we confirmed it. Reference pricing included all forms of price comparisons in which high/low dollar

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 70 (October 2006), 185–204

bune, *The Dallas Morning News*, *Los Angeles Times*, and *The New York Times*. Each advertisement was coded to identify whether reference pricing, limited-time availability, and sale announcements were included in the advertisement's headline or copy. Table 1 presents the frequency counts based on the content analysis. Several findings deserve mention. Approximately 29% of the advertisements featured reference pricing, limited-time availability, or a sale announcement in the headline or copy either alone or in combination. Sale announcements alone were displayed most often (9.0% of advertisements), followed by limited-time availability alone (7.9% of advertisements), and reference pricing alone (.9% of advertisements). All three appeared concurrently in 2.7% of the advertisements. Limited-time availability was present in 58.8% of the advertisements that featured reference pricing, and sale announcements appeared in 69.3% of those advertisements. Sale announcements appeared in 47.0% of the advertisements that contained limited-time offers. Finally, and most important, 87.2% of the 912 advertisements that displayed reference prices also included sale announcements or limited-time availability in the headline or copy. These findings demonstrate that retailers commonly pair reference price information with limited-time availability and sale announcements in their newspaper advertisements.

Our finding that reference pricing is rarely used in isolation in retail advertisement headlines and copy stands in stark contrast to behavioral research on reference price effects (for literature reviews, see Biswas, Wilson, and Licata 1993; Compeau and Grewal 1998). Prior behavioral reference-pricing research has focused almost exclusively on price comparisons and the manner in which they are portrayed but does not consider the coincidental use of limited-time availability and sale cues, which is the more common practice in retail advertisements.

The managerial significance of understanding why and when these variables influence consumer price and value perceptions becomes apparent in light of the dollar expenditures for retail newspaper advertisements that display them. Assuming that our content analysis of retail advertisements in four major newspapers generalize nationally, we estimated that $8.2 billion was spent in 2004 on retail advertising that contained one or more of these appeals, or approximately 18.3% of total retail newspaper advertising expenditures.[2]

Theoretical Perspectives and Research Purpose

The elaboration likelihood model (Petty and Cacioppo 1986) and the heuristic-systematic model (Chaiken 1980, 1987) are dual-process theories of persuasion that provide a foundation for this research. Both models postulate that high versus low motivation to process advertising information, operationalized through consumer involvement, results in different routes to persuasion: central/systematic (engaged through high-involvement processing) versus peripheral/heuristic (engaged through low-involvement processing). Different types of information are believed to have a persuasive impact under these different routes. This research examines the possibility that reference pricing, limited-time availability, and sale cues have inherent differences in their information-processing requirements, and thus consumers who are shopping (high involvement) versus those who are not shopping (low involvement) for a product featured in an advertisement respond to them differently. In particular, we argue that comprehending the full value implications inherent in reference price information

TABLE 1
Incidence of Reference Pricing, Limited-Time Offers, and Sale Announcements Featured in Newspaper Advertisements (N = 13,594)

	Reference Pricing Featured in Advertisement				
	Yes		No		
	Limited-Time Offer Featured in Advertisement		Limited-Time Offer Featured in Advertisement		
Sale Announcement	**Yes**	**No**	**Yes**	**No**	**Total**
Sale featured in advertisement	373 (2.7%)	259 (1.9%)	730 (5.4%)	1222 (9.0%)	2584 (19.0%)
Sale not featured in advertisement	163 (1.2%)	117 (.9%)	1079 (7.9%)	9651 (71.0%)	11,010 (81.0%)
Total	536 (3.9%)	376 (2.8%)	1809 (13.3%)	10,873 (80.0%)	13,594 (100%)

---

amounts were explicitly stated (e.g., "Regularly $_____, Now $_____"; "Manufacturer's List Price $_____, Sale Price $_____"). Limited-time availability included wording such as "_____ Days (Hours) Only"; "Hurry! Offer Ends Soon"; and "This Weekend Only!" Sale included the word "sale" and synonyms, such as "clearance/closeout," "liquidation," "going out of business," "deal," or "special" (with the word "sale" itself accounting for 80.1% of the occurrences).

[2] These estimates for 2004 assume zero ad revenue growth from $44.939 billion spent in 2003 (Newspaper Association of America 2006), with 70% of ad revenues coming from display advertisements (Belch and Belch 2004) and 90% of display revenues coming from advertisements that are 17 square inches or larger in size.

requires more thought and effort and, thus, greater involvement than limited time or sale cues.

The purpose of this investigation is to broaden the scope of reference price research to reflect contemporary retail advertising practices. Specifically, we examine the context in which and extent to which limited-time availability and sale cues in retail advertisements that feature reference pricing affect consumer price persuasion and store shopping intentions. We provide an overview of research on the major variables we examine, present our hypothesized expectations, and report our findings from three field experiments.

Reference Price Effects

A reference price advertisement is one in which a lower, current price is compared with a higher price either previously offered or offered by someone else. Considerable research in the marketing literature confirms the presence of reference price effects (Biswas, Wilson, and Licata 1993; Compeau and Grewal 1998; Kalyanaram and Winer 1995; Krishna et al. 2002; Mazumdar, Raj, and Sinha 2005). That is, when reference and retail (offering) prices are presented together, they are compared, and this comparison has a favorable effect on consumer deal perceptions relative to a condition in which reference prices are not presented. Different theoretical arguments have been advanced to account for reference price effects, including assimilation–contrast (Sherif, Sherif, and Nebergall 1965), adaptation-level theory (Helson 1964), and anchoring-adjustment principles (Tversky and Kahneman 1971). Although these theories make different predictions, they also share commonalities (see Mussweiler 2003). Each agrees that all judgments are relative and how a person evaluates something depends on what he or she uses as a basis for comparison. Furthermore, the location of an initial stimulus on an evaluative continuum (e.g., good versus bad and high versus low) can influence perceptions of a second stimulus on the same continuum, such that if the second is sufficiently different from the first, a person will judge it to be more different than it really is. In this way, a person may perceive a low price for a product as lower relative to a high reference price than when he or she evaluates the same low price in the absence of a reference price, all else being equal.

However, there is research that questions the magnitude and, indeed, the presence of reference price effects. This research suggests that for price- and promotion-sensitive consumers in particular, reference pricing effects may be confounded by other variables (Chang, Siddarth, and Weinberg 1999; Krishna, Currim, and Shoemaker 1991). In summarizing these and related studies, Neslin (2002) concludes that research is necessary to understand how other information contained in advertised promotions influences reference price effects. Our content analysis of retail newspaper advertisements confirms the need for such research.

Limited-Time Availability and Sale Cue Effects

With considerably less attention, the marketing literature has addressed limited-time availability or, more broadly, scarcity and sale cues in retail advertisements.[3] The evidence that exists suggests that limiting a consumer's opportunity to obtain a product makes the product more attractive and affects value perceptions. Empirical data in support of the effectiveness of scarcity in influencing value perceptions have been reported for fast food (Brannon and Brock 2001), batteries (Inman, Peter, and Raghubir 1997), nylon hosiery (Fromkin et al. 1971), and women's suits (Szybillo 1976). Anecdotal evidence of its effectiveness in selling other products also has been reported (see *The New York Times* 1999; *The Wall Street Journal* 2003).

With limited-time appeals in advertising, consumers are informed that they have a limited amount of time (e.g., "Three days only!") to act on an offering. Although there are disagreements about the mechanism by which scarcity cues, in general, affect value perceptions (see Lynn 1992), limited-time availability, in particular, seems to imply value because people have learned over the course of their lives to associate limited-time events with price deals (Cialdini 2001). Thus, favorable price inferences are readily accessible in memory on exposure to limited-time availability cues. Notably, research has documented the same response to sale cues that use this rationale (Inman and McAlister 1993; Inman, McAlister, and Hoyer 1990), and it has been suggested that retail advertisements that contain reference price information should include terms such as "sale" or "special" to communicate that the lower price or discount is only temporary (Grewal et al. 1998). Similarly, the need to include time limits in reference price research and to study their effect on consumer response to advertised price deals has been emphasized (Krishna et al. 2002). However, no research has systematically examined these cues in a reference price advertising context.

Consumer Involvement Effects

Following the work of Zaichkowsky (1985), we define "involvement" as the level of personal relevance that a product or purchase decision has for a consumer. Thus, we contend that people who are shopping for a product are more highly involved with information in an advertisement for that product than people who are shopping for something else. Consumer involvement has been shown to influence both the extent to which people process an advertisement and the type of information included (Chaiken 1980, 1987; Petty and Cacioppo 1986). Highly involved consumers are motivated to scrutinize information, such as prices, more fully, whereas less involved consumers are likely to apply simple heuristics, or judgment-relevant cues that are easily understood (Chen and Chaiken 1999), when processing information (Meyers-Levy and Peracchio 1996), such as the presence of limited-time availability and sale notifications in an advertisement.

From an information-processing perspective, we propose that people should make favorable price inferences more easily on the basis of limited-time availability and sale cues than reference price information in retail advertisements. We expect to find this because the implications of limited-time availability and sale announcements have been learned (Cialdini 2001) and can function as heuristic or peripheral cues (Chaiken 1980, 1987; Petty and Cacioppo 1986) that signal value. However, the value implied by reference pricing is not as easily inferred. Given the great variability that exists in the use of discounts in reference price advertisements (our content analysis revealed that offering prices ranged from 1% to 90% off advertised reference prices), consumers who want to have confidence in the accuracy of their judgments about price must *derive*, at some level, the value being offered rather than retrieve it from memory, as with limited-time availability and sale related cues. Thus, the cognitive processes involved in judging price favorability should be more complex when using reference price information than when using limited-time availability and sale cues. These differences in information-processing requirements should affect consumers when reference price information and limited-time availability or sale cues are most persuasively presented in retail advertisements.

---

[3] Our newspaper content analysis examined different scarcity variations. We based our decision to focus on limited time in this research on the finding that limited time not only was the most frequently used but also was used 3.5 times more often than the next most common variation (limited number/quantity).

Although multiple methods of conceptualizing and manipulating involvement have been used (for an overview, see Peter and Olson 2005), an operationalization that directly links active purchase intent among consumers with the processing of information common to retail advertisements in the marketplace is notably absent in the reference pricing literature. This is understandable because approximately 90% of the behavioral studies on reference pricing have been conducted in laboratory settings with student samples (Biswas, Wilson, and Licata 1993; Compeau and Grewal 1998). We contend that a clear stratification of personal product relevance is evident among actual consumers who are (are not) actively shopping for a particular good.[4] For example, a consumer who is engaged in shopping for a bedroom set will be more likely to process information in an advertisement for that product carefully than will a person who is interested in buying a sofa. A consumer who is shopping for a sofa may only incidentally or peripherally process information in an advertisement for a bedroom set. To our knowledge, no reference price advertising research has operationalized consumer involvement in a field setting and examined its influence on consumer price perceptions.

Study 1

Hypothesis

We designed Study 1 to examine the effects of reference price information and a limited-time availability cue on consumer price perceptions across levels of consumer involvement. We hypothesize that reference price information in an advertisement is likely to be effective in favorably influencing price perceptions only for consumers who are in the market for a product (high involvement). We expect that only these consumers will be motivated to expend the cognitive effort necessary to process reference price information fully. Conversely, we expect that consumers who are not in the market for a product (low involvement) will not have an interest in processing information to the degree necessary to make an informed judgment about the value represented by the reference price. This expectation is consistent with Yalch and Elmore-Yalch's (1984) suggestion that peripheral processors sometimes avoid numerical information. Thus, reference price information in advertisements should affect consumer price attitudes under high- but not low-involvement conditions.

We also hypothesize that limited-time availability in an advertisement should have a favorable effect on price judgments under both high- and low-involvement situations. As Petty and Wegener (1999) note, the same information can have an effect on persuasion across conditions of involvement, but different information-processing mechanisms are used. Specifically, although limited-time information can operate as a peripheral cue under low-involvement conditions, it would be evaluated "like all other available information" (Petty and Wegener 1999, p. 59) under high-involvement conditions. Its effect on price persuasion under high-involvement conditions would then depend on whether the limited-time availability of the offer has merit in evaluating price. Inman, Peter, and Raghubir (1997) argue that scarcity can operate both as a peripheral cue and as information evaluated on its merits. They find that consumers believe that limited-time availability has merit as an indicator of likely demand for the product. We believe that this should be especially true in times of consumer uncertainty, for example, when people consider the purchase of infrequently shopped goods, such as consumer durables. Under such conditions, and before the advertised product is physically examined, limited-time information should prove useful as a market valuation attribute, and therefore people should consider it along with other knowledge structures when evaluating price.

H_1: Limited-time availability significantly affects price attitudes among consumers who are not shopping for a product.

H_2: Limited-time availability and reference pricing significantly affect price attitudes among consumers who are shopping for a product.

These hypotheses are based on our expectation that consumers who are shopping (not shopping) for a product will (will not) allocate the cognitive effort necessary to process fully the value implications of reference prices in an advertisement for that product. Reference prices should affect price perceptions under high- but not under low-involvement conditions. However, the value implications of limited-time information should be easily cued from memory and used when consumers form price attitudes under both high- and low-involvement conditions. Thus, the type of ad information (limited-time availability, reference pricing) to which consumers are exposed will be differentially persuasive depending on the level of consumer shopping involvement.

---

[4] Involvement may be conceptualized as either "enduring" or "situational" (see Hoyer and MacInnis 2004). Enduring involvement reflects an inherent interest in a product independent of a purchase situation (see Chandrashekaran 2004; Chandrashekaran and Grewal 2003). Situational involvement reflects interest in a product caused by situational events, such as a purchase situation. We designed our manipulation of involvement to be situational: We wanted to evaluate ad effects on consumers who were actively engaged in marketplace shopping behavior.

Products and Prices

We tested our hypothesis over two different products: bedroom sets and sofas. Both products come in a wide range of prices and quality grades within and across furniture stores. Most people infrequently purchase bedroom sets and sofas and are inexperienced shoppers for these goods. Interviews with personnel at the store chain we used in this study, coupled with furniture industry research (Roy 2002), indicate that brand familiarity, store loyalty, and purchase frequency are low among furniture shoppers. We selected these products with the intent of heightening marketplace ambiguity for even high-involvement consumers. As Chandrashekaran (2004; see also Chandrashekaran and Grewal 2003) discusses, when consumers have a high degree of product and price knowledge, reference pricing may not work. Indeed, the effectiveness of advertised reference prices appears to depend on some level of consumer uncertainty about product price levels in the marketplace. We believe that such uncertainty will be true in general for infrequently purchased goods with low levels of brand familiarity.

Another reason for selecting bedroom sets and sofas is that they are often priced differently, thus allowing an examination of whether experimental results generalize across products with such prices. Store personnel and prospective customers (from stores in which study participants were obtained) were interviewed to determine the purchase price range for the advertised products and the content and layout of the advertisements. As a result of these interviews, we set reference prices at the upper end of the most common purchase price range, and we set the offering prices at the lower end of the most common purchase price range. We did this to enhance respondent perceptions of the lower prices as "deals" (Lichtenstein and Bearden 1989).[5] The offering prices for both products represent a 40% markdown from the reference price. Our ad review revealed that the mean markdown (off reference price) for furniture advertised in the area was 40.2%. Retail store personnel and prospective customers were also shown copies of the test advertisements. We confirmed that the reference prices were credible for the products being sold (see Compeau, Grewal, and Chandrashekaran 2002) and that the ad content and layout were representative of what was common to newspaper furniture advertisements in the area.

Procedure

Furniture shoppers were intercepted entering four different outlets of a well-known, moderately priced, regional furniture store chain in a large city in the Southwest. Bedroom sets and sofas constitute approximately 60% of the retailer's unit sales.

Shoppers were first asked to participate in a research study that was being jointly conducted by a local university and a (fictitious) local advertising association. Participants verbally responded to three questions, including what furniture they were shopping for that day. Those who answered either bedroom sets or sofas (but not both) were asked to participate in an advertising test. Participants were told they would examine print advertisements scheduled to be run the next day in an area newspaper. Those who agreed were handed a portfolio of seven professionally produced black-and-white print advertisements (8 × 10 inch reproductions of newspaper advertisements) and were verbally asked to look through them as they normally would.[6] After participants examined the advertisements, they filled out a questionnaire and were debriefed. No participant compensation was provided.

Experimental Manipulations and Design

We manipulated product involvement using the portfolio of seven advertisements to which participants were exposed. Six advertisements were filler advertisements for products or services unrelated to furniture. The fourth (target) advertisement was for a fictitious store (Furniture Direct) that sold either bedroom sets or sofas. Half of the consumers shopping for bedroom sets were given a portfolio with the advertisement for the bedroom set in the fourth position (high-involvement condition), and half were given the advertisement for the sofa in the fourth position (low-involvement condition). Likewise, half of the consumers shopping for sofas were given a portfolio with the sofa advertisement in the fourth position (high-involvement condition), and half were given the advertisement for the bedroom set in the fourth position (low-involvement condition). Sawyer and Howard (1991; see also Petty, Cacioppo, and Schumann 1983) use a similar involvement manipulation, in which different product advertisements alternatively served as the high- versus low-processing targets for people making a choice (versus not making a choice) for that product. This manipulation provides maximum design efficiency because it also allows a test of whether the results generalize across different products.

The two product test advertisements were similarly designed. The top (which stated the store name, Furniture Direct), subheadline (which stated, "We have more choices than anyone in town"), bottom (which provided a geographic address, a Web address, and a telephone number), and left-hand column (which provided financing-related information) of the advertisements were identical. However, one advertisement had a headline that asked, "Looking for a

---

[5] This pricing procedure also makes theoretical sense. By setting the advertised reference prices at the upper end of the most common purchase price range, we maximized the likelihood that the prices would influence consumers' internal reference prices, which are often conceptualized as expected or normal market prices or willingness to pay (see, e.g., Chandrashekaran 2004).

[6] We used black-and-white test advertisements on the basis of our ad review, which revealed that a majority (72.2%) of limited-time and reference price advertisements were black and white. This percentage was about the same for furniture advertisements. For consistency with the test advertisements, all control advertisements were also in black and white.

Bedroom Set?" and the other asked, "Looking for a Sofa?" Under the subheadline was either a photo of a five-piece bedroom set or photos of two sofas. Brief descriptions were provided for both products. Under the second sofa photo was the statement, "Take Your Pick—Either Sofa!" and under the bedroom set picture was the statement, "Five Piece Bedroom Set." There was additional information presented both under and above these statements that manipulated the presence of a "deal."

In the "no-deal" condition, the bedroom set was priced at "Now: $1725," and the sofas were priced at "Now: $734." In the limited-time condition, the words "Three Days Only!" were displayed above the statements "Five Piece Bedroom Set" or "Take Your Pick—Either Sofa." Again, the products were priced either "Now: $1725" (bedroom set) or "Now: $734" (sofas). In the reference price conditions, "Regularly $2876" (bedroom set) or "Regularly $1223" (sofas) was stated to the left of the current ("Now") prices. The combined condition was defined by the presence of both the reference price and the limited-time information (see Figure 1).[7]

We used a 2 (involvement: high versus low) × 4 (deal: no deal versus limited-time availability versus reference price versus combined) × 2 (product: bedroom set versus sofa) between-subjects design. A sample of 160 consumers participated in the experiment.

Measures

In the opinion book, participants were informed that different people were being asked questions about different advertisements from the portfolio. They were being asked for their opinions about the advertisement for the bedroom set or sofas. After opening the opinion book, participants were not permitted to review the advertisements again.

The dependent variable was consumer attitude toward the advertised product price.[8] Participants were asked, "I personally feel that the advertised selling price for the bedroom set/sofas is,..." and they responded on seven-point scales anchored by "good/bad," "undesirable/desirable," and "satisfactory/unsatisfactory." We performed a check on the involvement manipulation using two items. Participants were asked, "I found the advertisement for the bedroom set/sofas to be,..." and they responded on seven-point scales anchored by "personally involving/not personally involving" and "not relevant to me/relevant to me" (see Chaiken and Maheswaran 1994; Petty and Cacioppo 1979; Zaichkowsky 1985). All scale items were coded so that higher scores indicated more favorable price attitudes and greater involvement/relevance. Finally, we used three open-ended items to check for the possibility of hypothesis guessing and demand artifacts. These items revealed no problems in any of the experiments, and thus we do not discuss them further. We performed multiple comparisons using t-tests, with critical value adjustments based on the Bonferroni procedure (Rosenthal and Rosnow 1985; Rosenthal and Rubin 1984).

Manipulation Checks

We first performed a check of the involvement manipulation. Analysis of variance (ANOVA) results for both items revealed an involvement main effect. Consumers who were shopping for a bedroom set (sofa) rated advertisements for the respective products as more personally involving (M = 6.19; F(1, 144) = 179.47, $p < .0001$; $\eta^2 = .54$) and relevant (M = 5.65; F(1, 144) = 88.66, $p < .0001$; $\eta^2 = .36$) than those who were shopping for a sofa (bedroom set) (personally involving: M = 4.43; relevant to me: M = 4.36).

Price Persuasion Results

We found that the three items measuring attitude toward the advertised product price reliably measured the same underlying construct; we summed them ($\alpha = .90$). We found a significant main effect of deal (F(3, 144) = 21.95, $p < .0001$; $\eta^2 = .24$), in which price attitudes were less favorable in the no-deal condition than in any other condition. However, this was qualified by a significant interaction between involvement and deal (F(3, 144) = 16.53, $p < .0001$; $\eta^2 = .18$). Figure 2 plots the results of the involvement × deal interaction for Study 1.

Under low-involvement conditions, price attitudes were significantly higher for the combined treatment (M = 15.3) than for the reference price condition (M = 13.05; t = 2.63, $p < .05$). The mean for the limited-time condition (M = 16.0) was also more favorable than it was for the reference price treatment (t = 3.45, $p < .005$). Attitudes in the reference price treatment were identical to those in the no-deal condition (M = 13.05).

Under high-involvement conditions, the most favorable price attitudes were for the combined condition (M = 18.5), which were significantly higher (t = 2.04, $p < .05$, one-tailed test) than the reference price treatment (M = 16.85). Participants in the reference price condition perceived the advertised price more favorably (t = 4.15, $p < .0002$) than those who were exposed to the limited-time treatment (M = 13.3), who, in turn, judged the price more favorably (t = 2.22, $p < .05$, one-tailed test) than those who examined the advertisement with no deal (M = 11.5).

Discussion

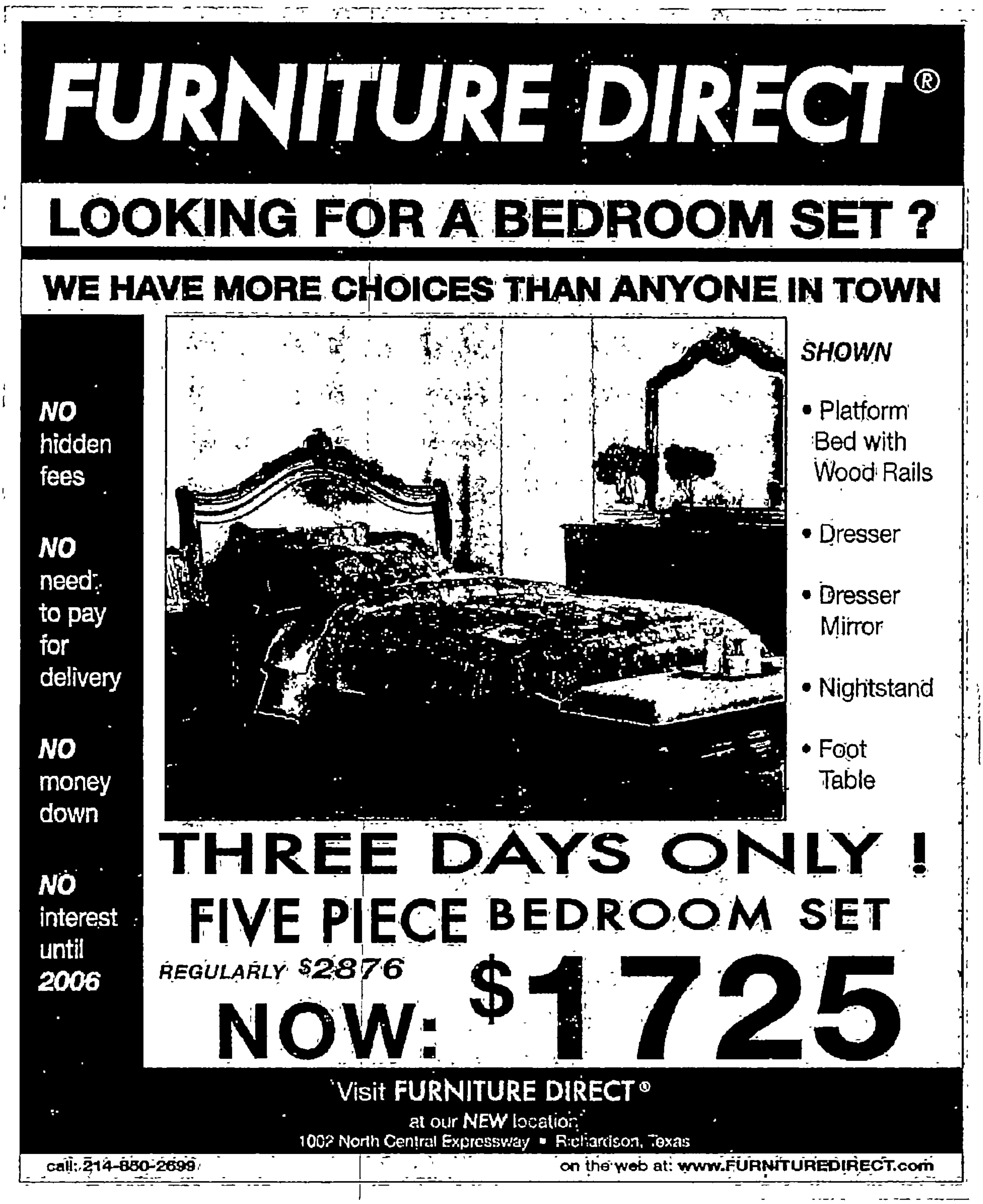
The results of Study 1 indicate that different deal information in advertisements is effective depending on whether consumers are or are not shopping for a particular product. When consumers are not shopping for a product, limited-time availability affects price perceptions, but reference pricing does not. Both the limited-time and the combined treatments resulted in more favorable price attitudes than the reference price condition, which did not differ significantly from the control (no-deal) condition. This is explained by viewing limited-time availability as a learned, value-oriented heuristic that is easily remembered and interpreted with respect to an advertised price. Limited time has utility in cuing value associations under low-involvement

---

[7] In all conditions across experiments, the advertised selling price ("Now") was presented. Only in the reference price and combined conditions was the higher reference price also included in the advertisements.

[8] This dependent variable is conceptually similar to "attitude toward the offer" used in previous studies (Bobinski, Cox, and Cox 1996; Lichtenstein and Bearden 1989).

FIGURE 1
Reference Price Advertisement with Limited-Time Appeal



FURNITURE DIRECT®

LOOKING FOR A BEDROOM SET ?

WE HAVE MORE CHOICES THAN ANYONE IN TOWN

NO hidden fees

NO need to pay for delivery

NO money down

NO interest until ***2006***

SHOWN

- Platform Bed with Wood Rails
- Dresser
- Dresser Mirror
- Nightstand
- Foot Table

THREE DAYS ONLY !

FIVE PIECE BEDROOM SET

REGULARLY $2876

NOW: $1725

Visit **FURNITURE DIRECT®**

at our ***NEW*** location

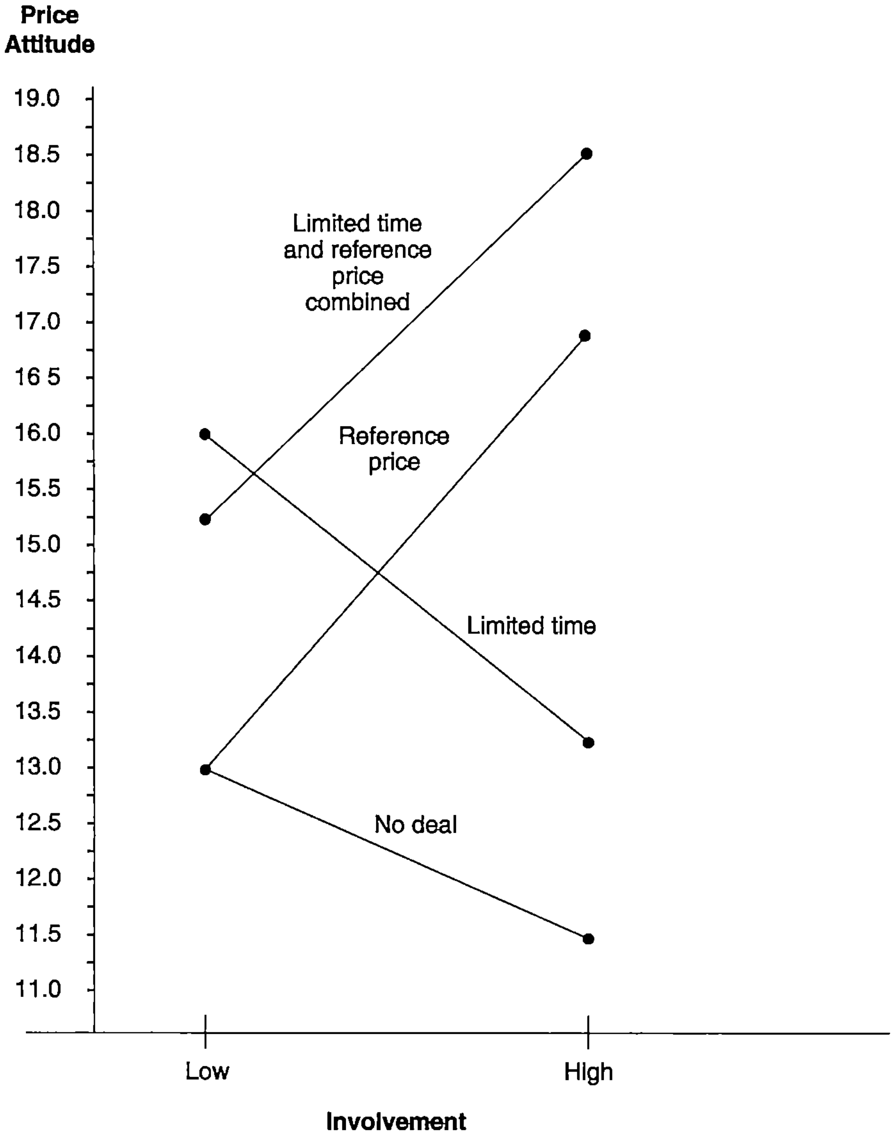
1002 North Central Expressway • Richardson, Texas

call: 214-850-2699

on the web at: www.FURNITUREDIRECT.com

conditions. However, the value implications inherent in reference pricing require a level of cognitive processing that exceeds what low-involvement consumers typically expend. The equality of the attitudinal means for the reference price and the no-deal conditions is consistent with the position that low-involvement consumers do not meaningfully process the reference price information. Importantly, the lack of a difference between the reference price and the no-

FIGURE 2
Involvement × Deal Interaction In Study 1



deal conditions further indicates that the mere presence of an advertised price reduction in the former treatment did not significantly affect price perceptions. Such a result might have occurred if simply the presence of a price reduction, in and of itself, served as an effective heuristic cue signaling a deal.

Under high-involvement conditions, consumer price attitudes in the combined condition were more favorable than the use of either reference prices or limited-time availability alone, which indicates that both factors jointly contributed to price persuasion. Given that the value indicated by the limited-time cue was directionally congruent with the reference prices, there was an additive effect on attitudes (also see Chaiken, Lieberman, and Eagly 1989). The finding that reference pricing had a stronger effect on attitudes than limited-time availability under high-involvement conditions indicates that consumers judged the former information as more salient or relevant than the latter when forming their attitudes toward the product prices. Such a judgment appears rational in that reference pricing provides information that directly (versus indirectly) addresses price. Thus, under high-involvement conditions, reference price information is more diagnostic of price value than limited-time information.

Finally, we did not obtain significant product effects (either main or interaction). This consistency in the results is notable given the different prices of the two furniture items. In summary, the results of Study 1 support the view that limited-time availability signals value when consumers are either shopping or not shopping for a product. Reference prices in advertisements are meaningfully processed only by consumers who are shopping for a particular product.

Study 2

Hypothesis

We conducted Study 2 to provide a further test of the position that the lack (presence) of respondent motivation to process reference prices in an advertisement accounts for the differential impact of that tactic in low- (high-) involvement conditions. To test this, we manipulated the use of "percentage-off" statements, which summarized the markdown size represented by the reference prices. For example, an advertisement that states, "Regularly $1223; Now: $734" would also state "40% Off" next to the prices. Although this practice is not common (our content analysis indicated that it is used in 20.0% of the advertisements that employ reference prices), it represents a potentially advantageous one. If the size of a markdown represented by the reference prices is salient to low-involvement consumers but requires too much effort for them to process, providing the percentage off in an advertisement should result in a significant increase in perceived favorability of the price. Conversely, for high-involvement consumers, who thoughtfully weigh and consider the information in an advertisement, providing a markdown summary would be unlikely to increase price perceptions favorably. These consumers are expected to derive the value implications of the price markdown on their own. We hypothesize that providing the size of a markdown in a reference price advertisement will increase price persuasion for low-involvement consumers but not for high-involvement consumers.

Study 2 investigates an additional dependent measure to broaden the domain of effects: consumer intentions to shop at the retail store that sponsored the advertisement (i.e., Furniture Direct). We added this element to test whether the independent variables are effective at bringing consumers into the market (shopping) for the advertised products, and it responds to the call to broaden reference price research by considering store patronage intentions (Mazumdar, Raj, and Sinha 2005). As a prerequisite for purchasing the goods, consumers must first be willing to shop at the store at which the products are sold. We anticipated the same effects for both the price persuasion and the store shopping intention measures. There are two bases for this expectation. First, given what is known about the attitude–behavioral intentions relationship (e.g., Fishbein and Ajzen 1975), theoretically, we expected a similar pattern of mean effects. Second, in the current experimental context, if consumers associate the price deal in the advertisement with the sponsoring retail store, favorable shopping intentions should result. Thus, we anticipate that the effects of involvement and deal on price attitudes and shopping intentions will be similar:

> H_3: Providing reference price markdown size favorably affects price attitudes and shopping intentions for low-involvement but not for high-involvement consumers.

We tested this hypothesis using an expanded version of the design implemented in Study 1.

Procedure, Manipulations, and Measures

We used the same procedure and involvement manipulation as in Study 1 with the same products (bedroom set and sofa). We also manipulated the deal variable as in Study 1, except that in certain conditions, the percentage off ("40% Off") was prominently stated (inside a shaded circle) to the left of the prices (see Figure 3). We restricted the conditions in which we used the percentage-off manipulations to those necessary for hypothesis testing. We used percentage-off statements in the reference pricing (called "reference price" versus "reference percentage off") and combined (called "combined" versus "combined percentage off") treatments.

We implemented a 2 (involvement: high versus low) × 6 (deal: no deal versus limited-time availability versus reference price versus reference percentage off versus combined versus combined percentage off) × 2 (product: bedroom set versus sofa) between-subjects design. A sample of 240 consumers participated in the experiment.

We used the price attitude items used in Study 1, followed by measures of store shopping intentions. Consumers were asked, "Think for a moment about the distance you traveled to get to this [retail furniture] store. If a Furniture Direct store were located about the same distance from where you live, how likely or unlikely would you be to shop at that Furniture Direct store?"[9] We assessed consumers' responses on two seven-point scales anchored by "very likely/very unlikely" and "not possible/very possible."

Manipulation Checks

As in Study 1, the ANOVA results for both manipulation check items revealed an involvement main effect. Consumers who were shopping for a bedroom set or sofa rated the respective advertisements for those products as more personally involving ($M = 6.36$; $F(1, 216) = 296.84$, $p < .0001$; $\eta^2 = .55$) and relevant ($M = 5.69$; $F(1, 216) = 88.58$, $p < .0001$; $\eta^2 = .28$) than those who were not shopping for these products (personally involving: $M = 4.38$; relevant to me: $M = 4.19$).

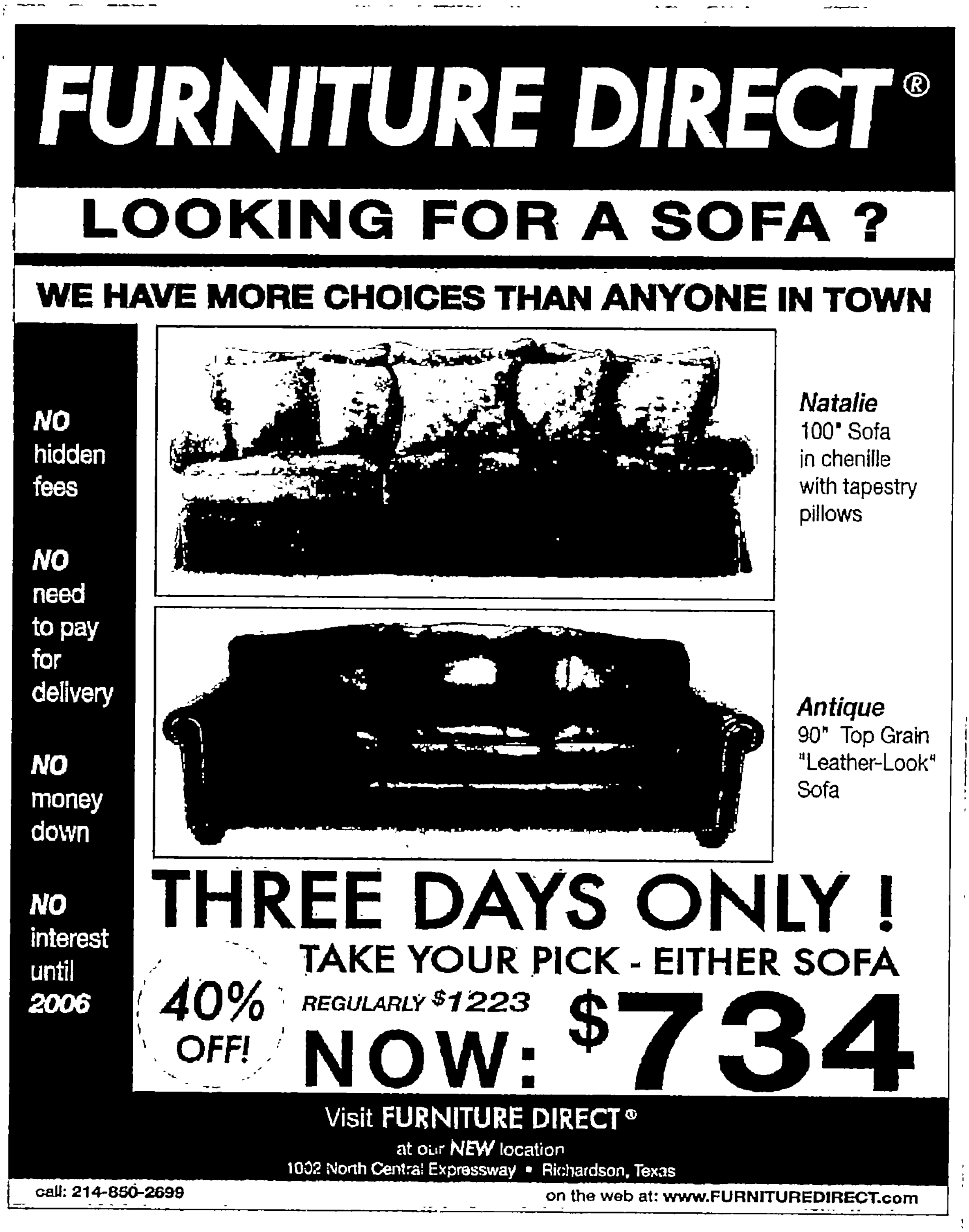
Price Persuasion Results

The three price attitude items reliably measured the same construct, and we summed them ($\alpha = .87$). The ANOVA results revealed a main effect of deal ($F(5, 216) = 14.06$, $p < .0001$; $\eta^2 = .21$), in which participants exposed to a deal were more favorable toward the offer price than those who were not exposed. Again, this was qualified by a significant interaction between involvement and deal ($F(5, 216) = 6.81$, $p < .0001$; $\eta^2 = .10$). Figure 4 shows the results of the involvement × deal interaction for Study 2.

---

[9] The wording of this item was necessary to control for distance effects on shopping intentions because the four stores used for shopper intercepts were located in different parts of the city and surrounding areas.

FIGURE 3
Reference Price Advertisement with Percentage-Off Summary and Limited-Time Appeal



FURNITURE DIRECT®

LOOKING FOR A SOFA ?

WE HAVE MORE CHOICES THAN ANYONE IN TOWN

NO hidden fees

NO need to pay for delivery

NO money down

NO interest until ***2006***

Natalie
100" Sofa
in chenille
with tapestry
pillows

Antique
90" Top Grain
"Leather-Look"
Sofa

THREE DAYS ONLY !

TAKE YOUR PICK - EITHER SOFA

40% OFF!

REGULARLY $1223

NOW: $734

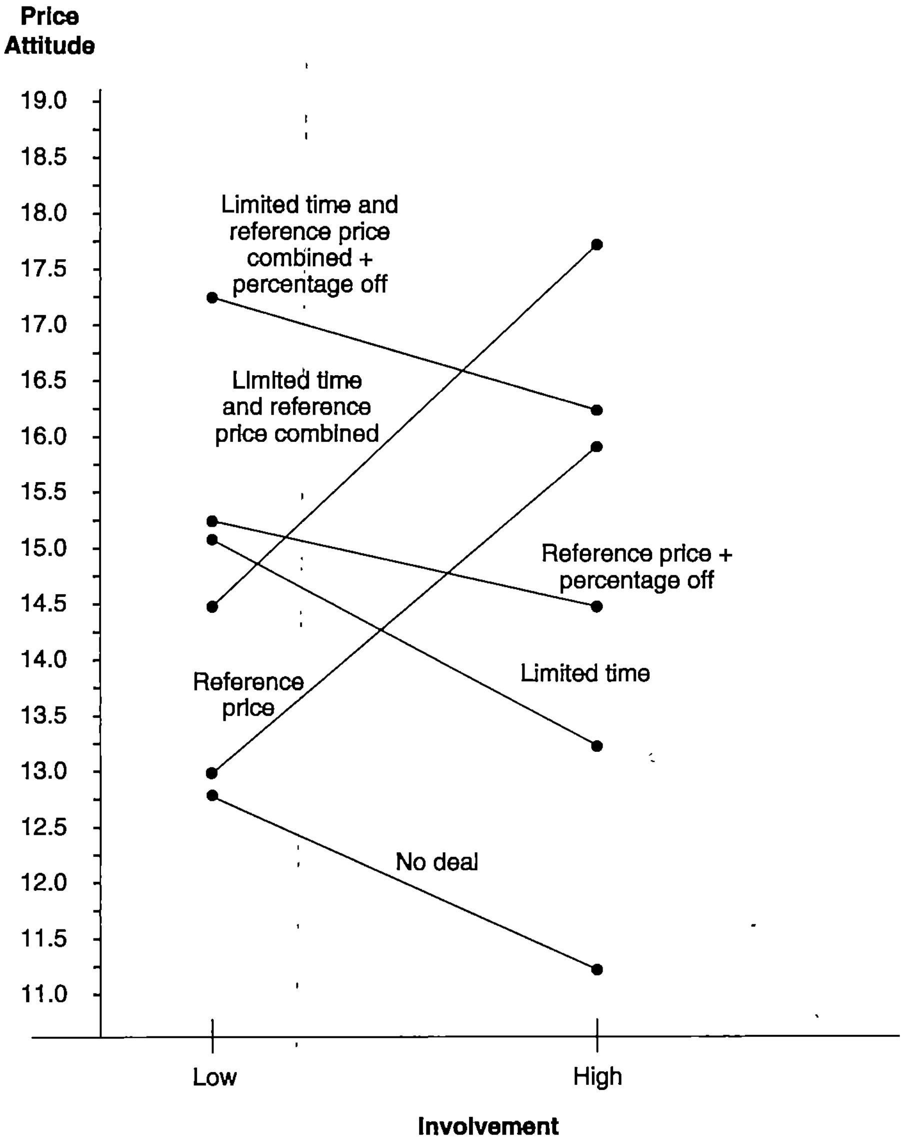
Visit **FURNITURE DIRECT®**
at our ***NEW*** location
1002 North Central Expressway • Richardson, Texas

call: 214-850-2699

on the web at: www.FURNITUREDIRECT.com

Under low-involvement conditions, the combined percentage-off group reported the most favorable price attitudes (M = 17.2); this was significantly higher (t = 3.10, $p < .005$) than the combined group (M = 14.45). Price attitudes for the reference percentage-off treatment (M = 15.2) were also higher (t = 2.53, $p < .05$) than the reference price group

FIGURE 4
Involvement × Deal Interaction in Study 2



(M = 12.95). The limited-time availability treatment (M = 15.1) resulted in more favorable price attitudes (t = 2.42, $p < .05$) than the reference price treatment, which did not significantly differ from the no-deal treatment (M = 12.8; t < 1). Across levels of the percentage-off manipulation, the combined treatment resulted in more favorable attitudes (M = 15.83) than the reference price treatment (M = 14.08; t = 2.79, $p < .05$).

Under high-involvement conditions, participants who were exposed to the combined treatment reported the most favorable price attitude (M = 17.8), though they were not reliably different (t = 1.63, $p > .10$) from those in the combined percentage-off condition (M = 16.35). The attitudes of those in the reference price condition (M = 15.95) were also not reliably different from those in the reference percentage-off treatment (M = 14.55; t = 1.58, $p > .10$), though the reference price treatment was more favorable than the limited-time condition (M = 13.25; t = 3.04, $p < .005$). Consumers in the limited-time condition had more favorable price attitudes than those who received the no-deal treatment (M = 11.25; t = 2.25, $p < .05$, one-tailed test). Finally, across levels of the percentage-off manipulation, the combined treatment resulted in more favorable price attitudes (M = 17.08) than the reference price condition (M = 15.25; t = 2.91, $p < .01$).

Store Shopping Intentions Results

The two items measuring store shopping intentions reliably measured the same construct ($\alpha = .85$); we recoded and summed them, with higher scores indicating greater store shopping intentions. The ANOVA results closely mirrored those we obtained with the price persuasion measure. There was a main effect of involvement (F(1, 216) = 11.02, $p <$

.001; $\eta^2 = .03$), such that high-involvement consumers were more likely to shop at Furniture Direct than low-involvement customers. There was also a main effect of deal ($F(5, 216) = 14.93$, $p < .0001$; $\eta^2 = .22$), such that participants exposed to an advertisement containing a deal were more likely to shop at Furniture Direct than those in the no-deal condition. However, these main effects were qualified by a significant involvement × deal interaction ($F(5, 216) = 5.91$, $p < .0001$; $\eta^2 = .09$).

Under low-involvement conditions, store shopping intentions for the combined percentage-off group ($M = 11.7$) were greater than those for the combined condition ($M = 9.45$; $t = 3.45$, $p < .003$). Similarly, the reference percentage-off group had a higher mean ($M = 10.9$) than the reference price condition ($M = 7.7$; $t = 4.91$, $p < .0001$). Participants who were exposed to a limited-time advertisement were more likely to indicate favorable store shopping intentions ($M = 9.75$) than those in the reference price treatment ($t = 3.15$, $p < .005$), which did not differ from the no-deal condition ($M = 7.8$; $t < 1$). Across levels of the percentage-off manipulation, the combined treatment had higher store shopping intentions ($M = 10.58$) than those in the reference price condition ($M = 9.3$; $t = 2.78$, $p < .01$).

Under high-involvement conditions, we found no significant difference between the combined percentage-off group ($M = 11.6$) and the combined condition ($M = 12.05$; $t < 1$). There was no difference between the reference percentage-off treatment ($M = 10.25$) and the reference price group ($M = 10.8$; $t < 1$). However, participants exposed to reference prices evidenced greater store shopping intentions than those in the limited-time treatment ($M = 9.5$; $t = 1.99$, $p < .05$, one-tailed test). The limited-time treatment resulted in marginally higher shopping intentions than for those in the no-deal condition ($M = 8.4$; $t = 1.69$, $p < .10$, one-tailed test). Summed over the percentage-off manipulation, the combined condition had greater store shopping intentions ($M = 11.83$) than the reference price treatment ($M = 10.53$; $t = 2.82$, $p < .01$).[10]

Discussion

Study 2 extends the findings we obtained in Study 1. The results indicate that summarizing reference price information in an advertisement has different effects depending on whether consumers are shopping for a product.

Under low-involvement conditions, attitudes and store shopping intentions for both the combined and the reference price treatments with the percentage-off statement were more favorable than the same respective treatments without the percentage-off statement. These findings are consistent with the position that providing summary information eases the processing requirements for comprehending the value implications of reference prices in advertisements. These results also replicate Chandrashekaran's (2004) and Chandrashekaran and Grewal's (2003) findings that reference prices with summary information can have a persuasive effect in low-involvement conditions. The lack of a significant difference in price attitudes and store shopping intentions for the low-involvement reference price and no-deal conditions indicates that without assistance, consumers who are not shopping for a product do not comprehend the value implications of reference prices; this is consistent with Study 1. However, the more favorable price attitudes and store shopping intentions for the combined and limited-time treatments than for the reference price condition indicate that limited-time availability has an effect on persuasion under low-involvement conditions.

A different pattern emerged under high-involvement conditions. When consumers were shopping for a product, we observed no reliable difference between exposure to either the combined or the reference price treatments and the same respective advertisements that included the percentage off. The price attitude and shopping intention means for the latter conditions were actually lower (less favorable) than those we obtained for treatments that did not summarize the markdown. Although the differences were not statistically significant at traditional levels, these results are conceptually congruent with the position that summarizing a benefit or conclusion for consumers who can be expected to process information actively on their own detracts from persuasion. Sawyer and Howard (1991) find that it is more persuasive for highly involved consumers to come to their own conclusions about the implications of a message. They argue that leaving an advertisement "open" (not providing a conclusion) causes central ad arguments to be processed more thoroughly, and resulting attitudes become more positive than in a condition in which conclusions are explicitly provided.

---

[10] At the request of a reviewer, we collected additional data to provide mediational evidence of the high-involvement reference pricing effect we observed in Studies 1 and 2. We compared prior search/price knowledge versus ad price processing explanations as potential mediators. We implemented a 2 (involvement: high versus low) × 2 (deal: reference price versus none) × 2 (product: bedroom set versus sofas) design using 80 furniture shoppers. We used the same dependent measures as in Studies 1 and 2 (price attitude: $\alpha = .83$; shopping intentions: $\alpha = .81$). We obtained two measures of prior search/price knowledge. Respondents were asked, "Within the past 30 days, how many stores have you shopped at that have sold bedroom sets (sofas)?" (coded 0–5) and "How knowledgeable or unknowledgeable are you regarding the prices of bedroom sets (sofas)?" (measured on scales anchored by "very knowledgeable/not knowledgeable"). We also collected two measures of ad price processing. Respondents were asked, "How much effort did you devote to thinking about the price information presented in the advertisement for the bedroom set (sofas)?" (measured on scales anchored by "high effort/low effort") and "I found the prices in the advertisement for the bedroom set (sofas) to be..." (measured on scales anchored by "not thought provoking/thought provoking"). All scaled items used seven-point scales. The two ad price processing items reliably measured the same underlying construct ($\alpha =$ 87); we summed them to form an ad price processing index. We first examined basic ANOVA effects. We found main effects of involvement and deal for both price attitude and store shopping intentions. More important, the critical involvement × deal interaction was also significant for both price attitude ($F(1, 72) = 18.57$, $p < .0001$; $\eta^2 = .16$) and store shopping intentions ($F(1, 72) = 10.02$, $p < .05$; $\eta^2 = .08$). The pattern of mean scores in these interactions mirrored previous findings (i.e., a differentiation between the two deal conditions for high-involvement but not for low-involvement respondents). We found main effects of involvement for the number of stores ($F(1, 72) = 16.28$, $p < .0001$; $\eta^2 = .18$), price knowledge ($F(1, 72) = 19.22$, $p < .0001$; $\eta^2 = .21$), and the ad price processing index ($F(1, 72) = 82.09$, $p < .0001$; $\eta^2 = .48$), with higher scores present in the high-involvement condition. We also found an involvement × deal interaction for the ad price processing index ($F(1, 72) = 6.56$, $p < .01$; $\eta^2 = .04$). We examined mediation of the interaction effect using Baron and Kenny's (1986) methodology. Neither the number of stores nor self-reported price knowledge was a significant or marginally significant mediator of the attitudinal or store shopping intention effects. However, the ad price processing index was a significant mediator of both (price attitude: $F(1, 71) = 6.66$, $p < .01$; $\eta^2 = .05$; store shopping intentions: $F(1, 71) = 5.16$, $p < .03$; $\eta^2 = .04$). The ad processing index partially mediated the deal × involvement interaction for price attitude ($F(1, 71) = 12.51$, $p < .001$; $\eta^2 = .10$) and store shopping intentions ($F(1, 71) = 6.05$, $p < .02$; $\eta^2 = .05$). These results support the position that scrutiny of price information presented in the advertisements by high-involvement respondents underlies the reference price effect we observed in Studies 1 and 2.

Consistent with Study 1, limited-time availability and reference pricing jointly contribute to price persuasion under high-involvement conditions in Study 2. The persuasive effect of limited-time availability under high-involvement conditions also replicates the Study 1 results. Finally, and also consistent with Study 1, we obtained no significant effects of the product variable.

Study 3

Hypothesis

We conducted Study 3 to examine informational and cue redundancy in a retail reference price advertisement under different involvement conditions. Despite the observation that multiple cues are common in many persuasion settings, limited research exists on their effectiveness (cf. Kalyanaraman and Oliver 2001). Study 3 provides insights into the operation of reference pricing and limited-time availability in retail advertisements by examining their effects when used with another value cue: the word "sale."

The suggestion that the word "sale" has acquired cue-related properties is not surprising given that its use is associated with the other persuasion techniques we examine herein. Our content analysis found that the word "sale" and related synonyms were used in 47.0% of advertisements with limited-time offers and in 69.3% of advertisements featuring reference pricing. Similar to the limited-time cue, we argue that the word "sale," through learned associations with advertised deals, has a favorable affect on price persuasion. Anderson and Simester's (2003) finding that simply displaying the word "sale" next to products in a store could increase sales up to 50% highlights its attention-grabbing ability, and it is consistent with the idea of automatic responding that typifies cue-related behaviors (Cialdini 2001).

Consistent with dual-process theories of persuasion (Chaiken 1987; Chen and Chaiken 1999; Petty and Cacioppo 1986), our contention has been that limited-time availability in retail advertisements operates as a peripheral or heuristic cue that facilitates persuasion under low-involvement conditions. However, under high-involvement conditions, limited-time availability is processed like other available information and is evaluated on its merits. If this is true, pairing the word "sale" with limited-time availability in an advertisement should result in no additional effect under high-involvement conditions, but it should contribute to persuasion under low-involvement conditions. We expect to find this because consumers who cogently process information in an advertisement would be more likely to deduce that limited-time prices are sales. Thus, when people examine an advertisement for a product that they are actively considering purchasing (i.e., under high-involvement conditions), the word "sale" is redundant information that should lack an effect, given the existence of a "Three days only!" price reduction.

However, redundancy between two cues should not impede the persuasive effects of either one for consumers who do not perceive them as redundant. We expect that recognizing the redundancy shared by "Three days only!" and the word "sale" will require a level of thinking that exceeds what is expended by consumers who view an advertisement for a product that does not match their current needs or interests (i.e., under low-involvement conditions). When advertisements try to communicate with people who are not shopping for a good, we believe that there is likely to be utility in stating what should be obvious because most consumers will not bother connecting the idea of the "Three days only!" message with that of the word "sale" unless the connection is made for them. In our view, leading consumers to a conclusion and then telling them what that conclusion is should have a persuasive impact in low-involvement contexts.

Finally, what would be the expectation when the word "sale" is used without being coupled with the limited-time treatment (i.e., when it is used alone with reference price information)? As a cue, it should add to persuasion under low-involvement conditions and should be evaluated on its merits under high-involvement conditions. However, we do not expect that the merits of a sale announcement under high-involvement conditions will have the same effect as announcing a limited-time event, given that the latter provides explicit parameters that define unavailability, whereas the former does not. In turn, the implied demand for a product associated with limited-time availability (Inman, Peter, and Raghubir 1997) and its corresponding market valuation utility should also be arguably stronger with limited-time availability than with sale announcements. Finally, the general idea of a sale from a consumer's perspective, though not the particulars of a limited-time event, is likely to be redundant with information logically obtained by considering reference price information under high-involvement conditions. That is, most consumers would interpret a price markdown as a sale.

In summary, we hypothesize that pairing the word "sale" with limited-time availability and reference prices in advertisements will increase price persuasion and store shopping intentions for low-involvement consumers but not for high-involvement consumers.

> H_4: Sale announcements in advertisements with limited-time availability and reference price appeals favorably affect price attitudes and store shopping intentions for low-involvement but not high-involvement consumers.

Procedure and Manipulations

We used the same procedure as in Studies 1 and 2. We manipulated involvement using the same products (bedroom set and sofa). We defined the deal variable with six treatment levels. The first three levels were the reference price, limited-time availability, and combined treatments, as we defined them in Studies 1 and 2. The next three levels were identical to the first three, except that the word "sale" was prominently highlighted (in a shaded circle) to the left of the pricing information in the advertisements (see Figure 5). We did not use the no-deal condition in Study 3 for cell size reasons.

We implemented a 2 (involvement: high versus low) × 6 (deal: limited time versus reference price versus combined versus limited-time sale versus reference price sale versus combined sale) × 2 (product: bedroom set versus sofa) between-subjects design. A sample of 240 consumers participated in the experiment. We used the same two dependent measures as in Study 2.

Manipulation Checks

The ANOVA results for both manipulation check items revealed a main effect of involvement. Participants who were shopping for a bedroom set or sofa rated the corresponding advertisements more personally involving (M = 5.87; F(1, 216) = 264.58, p < .00001; η^2 = .69) and relevant (M = 5.83; F(1, 216) = 153.13; p < .0001; η^2 = .57) than those who were shopping for the alternative product (sofa or bedroom set; personally involving: M = 4.28; relevant to me: M = 4.10).

Price Persuasion Results

Again, the three price persuasion items reliably measured the same construct, and we summed them to form an index (α = .83). The ANOVA results revealed main effects of involvement (F(1, 216) = 4.52, p < .04; η^2 = .01) and deal (F(5, 216) = 9.62, p < .0001; η^2 = .15), qualified by a significant interaction between them (F(2, 216) = 9.45, p < .0001; η^2 = .15).

Under low-involvement conditions, pairing reference prices with the word "sale" (M = 15.05) was more persuasive than the use of reference prices without the word "sale" (M = 12.25; t = 3.78, p < .0005). Similarly, pairing limited-time availability with the word "sale" was more persuasive (M = 17.1) than its use without the word "sale" (M = 15.2; t = 2.57, p < .005). Combining limited-time availability and reference price with the word "sale" was more persuasive (M = 16.6) than their use without the word "sale" (M = 14.95; t = 2.23, p < .05). Across levels of the sale variable, price attitudes for the combined condition (M = 15.78) were greater than those for the reference price treatment (M = 13.65; t = 4.10, p < .0002). Collapsed across levels of the sale variable, price attitudes in the limited-time availability condition (M = 16.15) were also more favorable than those in the reference price condition (t = 4.81, p < .0001).

Under high-involvement conditions, pairing limited time with the word "sale" did not result in more favorable price perceptions (M = 14.3) than its use without the word sale (M = 14.05; t < 1). Similarly, pairing reference prices with the word "sale" (M = 15.95) was not more favorable than its use without the word "sale" (M = 15.55; t < 1). The combined use of both techniques with the word "sale" (M = 17.7) was not more favorable than their use without the word "sale" (M = 17.45; t < 1). Across sale levels, the use of reference prices resulted in more favorable price attitudes (M = 15.75) than the use of limited-time availability (M = 14.18; t = 3.02, p < .05). Across sale levels, the combined reference price and limited-time condition resulted in more favorable price attitudes (M = 17.58) than the use of reference pricing by itself (M = 15.75; t = 2.47, p < .01). Figure 6 shows the results of the involvement × deal interaction.

Store Shopping Intentions Results

The two store shopping intentions items reliably measured the same underlying construct, and we summed them (α = .80). The results paralleled those for price persuasion. Main effects of involvement (F(1, 216) = 23.72, p < .0001; η^2 = .06) and deal (F(5, 216) = 18.26, p < .0001; η^2 = .23) were qualified by a significant interaction between those variables (F(5, 216) = 9.84, p < .0001; η^2 = .13).

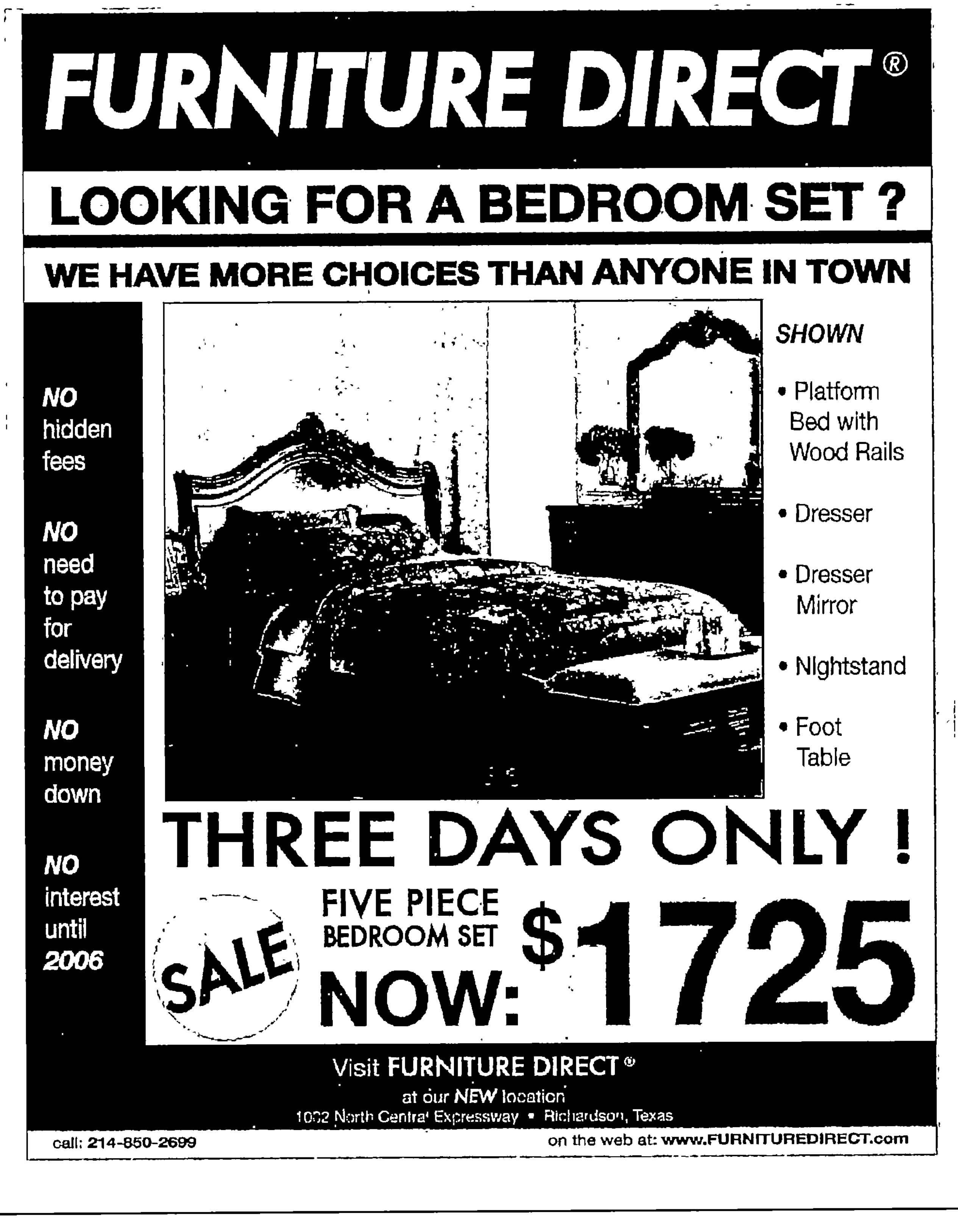
Under low-involvement conditions, when limited-time availability was advertised as a sale (M = 10.45), it was more persuasive than when it was advertised without the word "sale" (M = 8.75; t = 3.47, p < .0005). The same was true for the reference price (with sale = 8.6, without sale = 7.45; t = 2.35, p < .05) and combined (with sale = 11.0, without sale = 9.95; t = 2.14, p < .05) conditions. Collapsed across levels of sale, participants who were exposed to reference prices were less inclined to shop at Furniture Direct (M = 8.03) than those who were exposed to the limited-time treatment (M = 9.6; t = 4.53, p < .0001). Across sale levels, combining reference prices and limited-time availability resulted in greater store shopping intentions (M = 10.48) than reference prices by themselves (M = 8.03; t = 5.00, p < .000001).

Under high-involvement conditions, there was no significant difference in the persuasiveness of limited-time appeals when it was (M = 8.90) or was not (M = 8.95) advertised with the word "sale" (t < 1). The same results held for the reference price (with the word "sale" = 10.3, without the word "sale" = 10.45; t < 1) and the combined (with the word "sale" = 12.05, without the word "sale" = 11.40; t = 1.33, p > .10) treatments. Across sale levels, reference pricing was more persuasive (M = 10.38) than limited-time availability (M = 8.93; t = 4.14, p < .0001) on store shopping intentions. Across sale levels, the combined condition resulted in greater store shopping intentions (M = 11.73) than the reference price condition (M = 10.38; t = 2.69, p < .01).

Discussion

Study 3's results demonstrate that under low-involvement conditions, advertised announcements of a sale and a limited time to take advantage of a lower price independently contribute to favorable consumer price perceptions and higher store shopping intentions. This can be explained by noting that though sales, by definition, are supposed to be limited-time events, when consumers are perusing advertisements without personal interest, they do not think at a

FIGURE 5
Limited-Time Sale Advertisement



FURNITURE DIRECT®

LOOKING FOR A BEDROOM SET ?

WE HAVE MORE CHOICES THAN ANYONE IN TOWN

NO hidden fees

NO need to pay for delivery

NO money down

NO interest until ***2006***

SHOWN

- Platform Bed with Wood Rails
- Dresser
- Dresser Mirror
- Nightstand
- Foot Table

THREE DAYS ONLY !

SALE

FIVE PIECE
BEDROOM SET

NOW: $1725

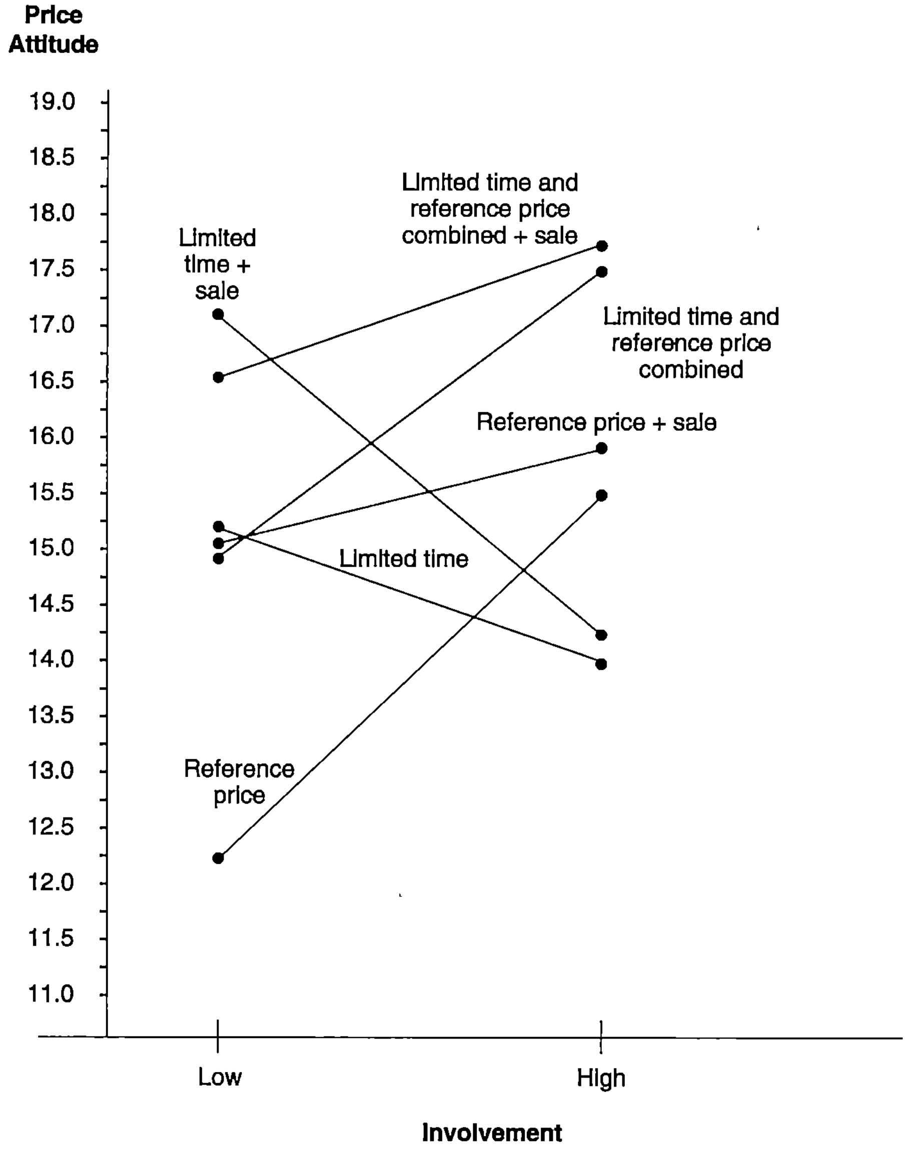
Visit **FURNITURE DIRECT®**
at our *NEW* location
1002 North Central Expressway • Richardson, Texas

call: **214-850-2699**

on the web at: **www.FURNITUREDIRECT.com**

level deep enough to make the connection. Thus, value cues that are otherwise logically redundant can independently contribute to persuasion under low-involvement conditions. Under high-involvement conditions, however, the value implications of the word "sale" are no different from those of a limited-time event (or reference prices). These findings

FIGURE 6
Involvement × Deal Interaction in Study 3



do not suggest that the word "sale" has no utility under high-involvement conditions. Rather, when consumers cogently process ad-related information, sale announcements contribute nothing unique to price perceptions or shopping intentions that these consumers do not already attain when they consider limited-time availability and reference price information. Finally, this study replicates Studies 1 and 2 in terms of effect (limited-time availability and reference pricing) differentiation across levels of consumer involvement and consistency of the results across products.[11]

General Discussion

Are limited-time availability and sale cues effective in influencing consumer attitudes toward product prices in retail newspaper advertisements that feature reference pricing? Can they also draw people into stores? If it is assumed that retailer advertising practices represent what "works" in the marketplace, the answer is yes according to our content analysis of 13,594 retail advertisements. This analysis revealed that 87.2% of retail advertisements featuring reference pricing also included limited-time availability or sale announcements in the headline or copy. However, the results of three field experiments across two different products consistently demonstrate significant variability in the effectiveness of these techniques, depending on whether consumers are or are not shopping for an advertised product.

---

[11] As a reviewer requested, we gathered additional data to examine whether low-involvement furniture shoppers versus nonfurniture shoppers would respond differently to the limited-time + sale appeals. Using the same method we previously discussed, we exposed low-involvement furniture (N = 40) and nonfurniture (N = 40; electronics products) shoppers to either the bedroom set or the sofa advertisements across the limited-time + sale versus the no-deal conditions. We implemented a 2 (shopper: furniture versus electronics) × 2 (deal: limited time + sale versus no deal) × 2 (product: bedroom set versus sofa) between-subjects design. We examined price attitude ($\alpha = .85$) and store shopping intentions ($\alpha = .80$) Only the effect of deal was significant for attitude ($F(1, 72) = 35.24$, $p < .0001$; $\eta^2 = .33$) and store shopping intentions ($F(1, 72) = 19.18$, $p < .0001$; $\eta^2 = .21$), and the limited-time + sale advertisement resulted in more favorable price attitudes and shopping intentions at the designated store than exposure to the no-deal condition. The lack of interaction effects indicates that this finding is statistically consistent for furniture and nonfurniture shoppers across the advertised products.

Substantive Findings

Using dual-process theories of persuasion (the elaboration likelihood model and the heuristic-systematic model) as a foundation, we have specified the conditions under which these techniques result in more favorable price attitudes and store shopping intentions and when they do not. Under low-involvement conditions (when consumers are not actively shopping for a product), a limited-time price reduction or a sale announcement affects price perceptions and shopping intentions, but the size of the reduction, represented by reference prices, does not. Under high-involvement conditions (when consumers are actively shopping for a product), both reference prices and limited-time information favorably affect price perceptions and shopping intentions, though additional explanatory power is not achieved through the use of sale announcements.

The reference price results can be understood by recognizing that comprehension of the value implications of those prices requires a level of processing effort that exceeds (does not exceed) the level that people typically expend when viewing advertisements for products that do not have (have) immediate relevance to them. Mediational findings support the position that more diligent processing of prices presented in the advertisements, as opposed to prior store search or product price knowledge, explains the favorable effect of reference pricing under high-involvement conditions. The effects of limited-time availability and sale announcements can be understood in terms of their learned heuristic value under low-involvement conditions and their informational merit under-high involvement conditions. Under high-involvement conditions, the implications of a sale are redundant with the information derived through a careful consideration of limited-time availability and reference price information. Under low-involvement conditions, however, consumers do not perceive the redundancy between limited-time availability and sale announcements, and therefore both cues affect persuasion. Thus, in general, we argue that reference pricing, limited-time availability, and sale cues have inherent differences in terms of their information-processing requirements and outcomes. Overall, the three experiments help predict the effects of jointly using different value-oriented persuasion techniques with reference price information in retail advertisements.

Notably, not only do our findings demonstrate strength and consistency, but they are also the result of field experiments. Intercepts for all three studies used engaged shoppers who had other objectives in mind when they volunteered to participate in the studies (without compensation). The outdoor environment in which testing took place was complete with noise and other distractions. Yet, despite these interferences, the effects were reliable across multiple studies. We suspect that one reason for the strength and consistency of the effects was the nature of the involvement manipulation. We believe that intercepting consumers in the process of shopping provides a stronger manipulation of involvement as personal product relevance than is typically found in laboratory manipulations of the construct. The average variance (across experiments) explained (η^2) in the two manipulation check measures was 49.8%. The average variance explained for the critical involvement × deal interaction was 14.8% for price persuasion and 10.0% for store shopping intentions. These effect sizes are about two to three times the average sizes that Peterson, Albaum, and Beltramini (1985) report in their review of effect sizes.

Managerial Perspectives

When used by itself, reference price advertising results in favorable price perceptions and store shopping intentions only among consumers who are shopping for a particular product. Reference pricing is an effective persuasion strategy among consumers who are of the greatest interest to retailers, namely, those with an immediate product interest. Consumers who lack product involvement do not process the implications of reference prices in retail advertisements to the degree necessary to grasp the value implications. However, managers might still consider using reference pricing with a markdown summary statement to appeal to a broader audience. We found that a markdown summarization results in increased price persuasion and shopping intentions for low-involvement shoppers but does not significantly detract from persuasion for high-involvement shoppers. The value in marketing to consumers without an immediate product interest lies in the likelihood of future rather than immediate sales being facilitated by associating the retailer with favorable pricing practices in the minds of consumers.

Reference pricing paired with limited-time availability results in more favorable price perceptions and store shopping intentions than the use of either technique alone among consumers who are shopping for a product. Practitioner insight regarding this point is shown by the finding that reference pricing and limited-time availability are used together in retail advertisements more frequently than reference pricing is used alone. Importantly, however, the three studies we report herein qualify this insight by providing consistent evidence across two products that both techniques independently contribute to price persuasion and shopping intentions only when consumers are actively shopping for a product. Nevertheless, as a general statement, the joint use of reference pricing and limited-time appeals might still be advisable because the effects of the two techniques appear to compensate for each other under different conditions of involvement. Specifically, reference price effects are (are not) significant when consumers are (are not) shopping for a product, but significant limited-time effects are weaker (stronger) under those same conditions.

Sale announcements in reference price advertisements appear to have no drawbacks. In developing this study, discussions with retailers revealed a growing reluctance to use the term "sale" in advertisements because of a perception that it is overused. Markdowns in retail advertising are commonplace, and consumers often discount them (Bobinski, Cox, and Cox 1996). Our content analysis indicates that the word "sale" is used with moderate frequency. In our field studies, however, a sale announcement (when used in a reference price advertisement) did not detract from persuasion when consumers were shopping for a product, but it contributed to persuasion when they were not. The potential utility of sale cues being presented to disinterested consumers should not be underestimated, especially because those cues do not appear to lower evaluations of consumers who are currently in the market. For example, sale cues might have the potential to convert nonshoppers into shoppers.

Directions for Further Research

Further research needs to examine conditions other than involvement that have the potential to moderate the effects of reference pricing. Given their potential influence on consumer motivation, two constructs that might affect the processing of reference price information are the level of prior product knowledge (Alba and Hutchinson 1987) and deal proneness (Krishna 1994). Research should also explore variables that have the potential to magnify reference price effects. The joint effect of reference prices and limited-time availability under high-involvement conditions suggests that when consumers are shopping for a product, the value implications of reference prices are accentuated when they are limited in time. The utility of different reference price formats in converting nonshoppers into shoppers also needs to be examined.

Testing whether our limited-time findings extend to other scarcity variations would develop the generality of our results. For example, it is an empirical question whether respondents would be equally sensitive to both limited-time and limited-quantity manipulations across involvement conditions. Thus, appropriate caution needs to be taken in terms of generalizing our results across different types of scarcity because we examine only limited-time in these studies.

Further research needs to consider different methodologies for investigating reference price advertising effects. Prior research has typically employed forced exposure to print advertisements and a passive manipulation of involvement within the context of laboratory studies using student participants. The current research employed free exposure to print advertisements and an active manipulation of involvement in a field setting in which consumers were engaged in shopping behavior. Although both methodologies have merit, we selected the latter because of its closer correspondence to the processing realities of consumers in the marketplace, as Wells (1993) and Kover (1995) advocate. Nevertheless, we recognize that our manipulation of involvement also has potential limitations. For example, by intercepting shoppers, the study samples were composed of goal-directed consumers. Some commentators might argue that low-involvement consumers may have been overly dismissive of product advertisements that did not reflect their current interests. In the leisurely environment of a home, low-involvement processors might pay more attention to such advertisements than they did in our studies. We acknowledge that there are different ways to manipulate involvement, and the nature of that manipulation may manifest itself in the results obtained.

The possibility that demand characteristics explain the results appears remote given the nature of the interaction effects we obtained. Specifically, a demand explanation for our findings can account for neither the low-involvement results nor the pattern of effects observed for the high-involvement condition. Furthermore, there was no evidence of hypothesis guessing or compliant participant behavior in answers to postexperimental questions. The results varied consistently with theoretical expectations, which is also inconsistent with a demand explanation for the results.

Although much of modern advertising uses multiple persuasive techniques in the same advertisement, academic research has often focused on examining individual techniques in isolation. A particular value of our studies is that they consider how different techniques are used in advertisements that feature reference pricing and how consumers should theoretically respond to them and then test the corresponding predictions in a field setting. Further research on reference price advertising should be expanded to explore the use of multiple persuasion tools commonly employed in retail advertising practice. We hope that our efforts will provide a framework for future investigations that broaden the scope of reference price advertising research.

REFERENCES

Alba, Joseph A. and Wesley Hutchinson (1987), "Dimensions of Consumer Expertise," *Journal of Consumer Research*, 13 (March), 411–54.

Anderson, Eric T. and Duncan I. Simester (1998), "The Role of Sale Signs," *Marketing Science*, 17 (2), 139–55.

——— and ——— (2003), "Mind Your Pricing Cues," *Harvard Business Review*, 81 (September), 96–103.

Baron, Robert M. and David A. Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (December), 1173–82.

Belch, George E. and Michael A. Belch (2004), *Advertising and Promotion: An Integrated Marketing Communications Perspective*. Burr Ridge, IL: McGraw-Hill/Richard D. Irwin.

Biswas, Abhijit, Elizabeth J. Wilson, and Jane W. Licata (1993), "Reference Pricing Studies in Marketing: A Synthesis of Research Results," *Journal of Business Research*, 27 (July), 239–56.

Bobinski, George S., Jr., Dena Cox, and Anthony Cox (1996), "Retail 'Sale' Advertising, Perceived Retailer Credibility, and Price Rationale," *Journal of Retailing*, 72 (3), 291–306.

Brannon, Laura A. and Timothy C. Brock (2001), "Limiting Time for Responding Enhances Behavior Corresponding to the Merits of Compliance Appeals: Refutations of Heuristic-Cue Theory in Service and Consumer Settings," *Journal of Consumer Psychology*, 10 (3), 135–46.

Chaiken, Shelly (1980), "Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion," *Journal of Personality and Social Psychology*, 39 (5), 752–66.

——— (1987), "The Heuristic Model of Persuasion," in *Social Influence: The Ontario Symposium*, Vol. 5, M.P. Zanna, J.M. Olson, and C.P. Herman, eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 472–89.

———, A. Lieberman, and Alice H. Eagly (1989), "Heuristic and Systematic Information Processing Within and Beyond the Persuasion Context," in *Unintended Thought*, J.S. Uleman and J.A. Bargh, eds. New York: Guilford Press, 212–52.

——— and Durairaj Maheswaran (1994), "Heuristic Processing Can Bias Systematic Processing: Effects of Source Credibility, Argument Ambiguity, and Task Importance on Attitude Judgment," *Journal of Personality and Social Psychology*, 66 (3), 460–73.

Chandrashekaran, Rajesh (2004), "The Influence of Redundant Comparison Prices and Other Price Presentation Formats on Consumers' Evaluations and Purchase Intentions," *Journal of Retailing*, 80 (1), 53–66.

——— and Dhruv Grewal (2003), "Assimilation of Advertised Reference Prices: The Moderating Role of Involvement," *Journal of Retailing*, 79 (1), 53–62.

———, Kent Monroe, and Madhubalan Viswanathan (2004), "Size-Value Incongruence in Comparative Price Advertisements," working paper, Department of Marketing, College of Business Administration, Fairleigh Dickinson University.

Chang, Kwangpil, S. Siddarth, and Charles B. Weinberg (1999), "The Impact of Heterogeneity in Purchase Timing and Price Responsiveness on Estimates of Sticker Shock Effects," *Marketing Science*, 18 (Spring), 178–92

Chen, Serena and Shelly Chaiken (1999), "The Heuristic-Systematic Model in its Broader Context," in *Dual Process Theories in Social Psychology*, Shelly Chaiken and Yaacov Trope, eds. New York: Guilford Press, 73–96.

Cialdini, Robert B. (2001), *Influence: Science and Practice*. Boston: Allyn & Bacon.

Compeau, Larry D. and Dhruv Grewal (1998), "Comparative Price Advertising: An Integrative Review," *Journal of Public Policy & Marketing*, 17 (Fall), 257–73.

———, ———, and Rajesh Chandrashekaran (2002), "Comparative Price Advertising. Believe It or Not," *Journal of Consumer Affairs*, 36 (2), 284–94.

Fishbein, Martin and Icek Ajzen (1975), *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.

Fromkin, H.L., J.C. Olson, R.L. Dipboye, and D.A. Barnaby (1971), "A Commodity Theory Analysis of Consumer Preferences for Scarce Products," paper presented at the American Psychological Association meeting, Washington, DC (September 3–7).

Grewal, Dhruv, R. Krishna, Julie Baker, and Norm Borin (1998), "The Effect of Store Name, Brand Name and Price Discounts on Consumers' Evaluations and Purchase Intentions," *Journal of Retailing*, 74 (4), 331–52.

Helson, H. (1964), *Adaptation Level Theory* New York: Harper and Row.

Hoyer, Wayne D. and Deborah J. MacInnis (2004), *Consumer Behavior*. New York: Houghton Mifflin.

Inman, J. Jeffrey and Leigh McAlister (1993), "A Retailer Promotion Policy Model Considering Promotion Signal Sensitivity," *Marketing Science*, 12 (Fall), 339–55.

———, ———, and Wayne D. Hoyer (1990), "Promotion Signal: Proxy for a Price Cut?" *Journal of Consumer Research*, 17 (June), 74–81.

——— Anil C. Peter, and Priya Raghubir (1997), "Framing the Deal: The Role of Restrictions in Accentuating Deal Value," *Journal of Consumer Research*, 24 (June), 68–79.

Kalyanaram, Gurumurthy and Russell S. Winer (1995), "Empirical Generalizations from Reference Price Research," *Marketing Science*, 14 (8), G161–69.

Kalyanaraman, Sriram and Mary Beth Oliver (2001), "Information Processing of Multiple Peripheral Cues," paper presented to the Information Systems Division at the 51st annual conference of the International Communication Association, Washington, DC (May 24–28).

Kover, Arthur J. (1995), "Copywriters' Implicit Theories of Communication: An Exploration," *Journal of Consumer Research*, 21 (March), 596–611.

Krishna, Aradna (1994), "The Effect of Deal Knowledge on Consumer Purchase Behavior," *Journal of Marketing Research*, 31 (February), 76–91.

———, Richard Briesch, Donald R. Lehmann, and Hong Yuan (2002), "A Meta-Analysis of the Impact of Price Presentation on Perceived Savings," *Journal of Retailing*, 78 (2), 101–118.

———, Imran S Currim, and Robert W Shoemaker (1991), "Consumer Perceptions of Promotional Activity," *Journal of Marketing*, 55 (April), 4–16.

Lichtenstein, Donald R. and William O. Bearden (1989), "Contextual Influences on Perceptions and Merchant-Supplied Reference Prices," *Journal of Consumer Research*, 16 (June), 55–66.

Lynn, Michael (1992), "The Psychology of Unavailability," *Basic and Applied Social Psychology*, 13 (1), 3–7.

Mazumdar, Tridib, S.P. Raj, and Indrajit Sinha (2005), "Reference Price Research: Review and Propositions," *Journal of Marketing*, 69 (October), 84–102.

Meyers-Levy, Joan and Laura Peracchio (1996), "Moderators of the Impact of Self-Reference on Persuasion," *Journal of Consumer Research*, 22 (March), 408–423.

Mussweiler, Thomas (2003), "Comparison Processes in Social Judgment: Mechanisms and Consequences," *Psychological Review*, 110 (3), 472–89.

Neslin, Scott A. (2002), *Sales Promotion*. Cambridge, MA: Marketing Science Institute.

Newspaper Association of America (2006), "Facts About Newspapers 2004," (January 4), (accessed June 27, 2006), [available at http://www.naa.org/Trends-and-Numbers/Market-Databank/Industry-Statistics.aspx].

The New York Times (1999), "Scarcity Makes the Heart Grow Fonder," (September 5), Section 9, 1.

Peter, J. Paul and Jerry C Olson (2005), *Consumer Behavior and Marketing Strategy*. Burr Ridge, IL: McGraw-Hill.

Peterson, Robert A., Gerald Albaum, and Richard F. Beltramini (1985), "A Meta-Analysis of Effect Sizes in Consumer Behavior Experiments," *Journal of Consumer Research*, 12 (June), 97–103.

Petty, Richard E. and John T. Cacioppo (1979), "Issue Involvement Can Increase or Decrease Persuasion by Enhancing Message-Relevant Cognitive Responses," *Journal of Personality and Social Psychology*, 37 (10), 1915–26.

——— and ——— (1986), *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

———, ———, and David Schumann (1983), "Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement," *Journal of Consumer Research*, 10 (2), 135–45.

——— and Duane T. Wegener (1999), "The Elaboration Likelihood Model: Current Status and Controversies," in *Dual Process Theories in Social Psychology*, Shelly Chaiken and Yaacov Trope, eds. New York: Guilford Press.

Rosenthal, Robert and Ralph A. Rosnow (1985), *Contrast Analysis: Focused Comparisons in the Analysis of Variance*. Cambridge: Cambridge University Press.

——— and D.B. Rubin (1984), "Multiple Contrasts and Ordered Bonferroni Procedures," *Journal of Educational Psychology*, 76 (6), 1028–1034.

Roy, Joella (2002), "The Consumer Purchasing Process–Part 2," *Furniture World*, (October), (accessed December 4, 2004), [available at http//:www.furnitureworld@furninfo.com].

Sawyer, Alan G. and Daniel J. Howard (1991), "Effects of Omitting Conclusions in Advertisements to Involved and Uninvolved Audiences," *Journal of Marketing Research*, 28 (November), 467–74.

Sherif, C., M. Sherif, and R.E. Nebergall (1965), *Attitude and Attitude Change*. Philadelphia: Saunders.

Szybillo, G.L. (1976), "A Situational Influence on the Relationship of a Consumer Attribute to New Product Attractiveness," *Journal of Applied Psychology*, 60 (October), 652–55.

Tversky, A. and D. Kahneman (1971), "Judgment Under Uncertainty: Heuristics and Biases," *Science*, 185 (September), 1124–31.

The Wall Street Journal (2003), "The T-Shirt You Can't Get," (October 10), W1.

Wells, William D. (1993), "Discovery-Oriented Consumer Research," *Journal of Consumer Research*, 19 (March), 489–504.

Yalch, Richard F. and Rebecca Elmore-Yalch (1984), "The Effect of Numbers on the Route to Persuasion," *Journal of Consumer Research*, 11 (June), 522–27.

Zaichkowsky, Judith L. (1985), "Measuring the Involvement Construct," *Journal of Consumer Research*, 12 (December), 341–52.

JOURNAL OF MARKETING
SUBJECT AND AUTHOR INDEX TO VOLUME 70
2006

SUBJECT INDEX

Item Number — *Authors*

ADVERTISING

1. Broadening the Scope of Reference Price Advertising Research: A Field Study of Consumer Shopping Involvement
Vol 70, No. 4, October 2006, 185–204
Daniel J. Howard & Roger A. Kerin

2. Crossover Dreams: Consumer Responses to Ethnic-Oriented Products
Vol. 70, No. 2, April 2006, 35–51
Sonya A. Grier, Anne M. Brumbaugh, & Corliss G. Thornton

3. It Just Feels Good: Customers' Affective Response to Touch and Its Influence on Persuasion
Vol. 70, No. 4, October 2006, 56–69
Joann Peck & Jennifer Wiggins

4. Marketing's Credibility: A Longitudinal Investigation of Marketing Communication Productivity and Shareholder Value
Vol. 70, No. 4, October 2006, 70–91
Xueming Luo & Naveen Donthu

5. Scents and Sensibility. When Do (In)Congruent Ambient Scents Influence Product Evaluations?
Vol. 70, No. 3, July 2006, 32–43
Anick Bosmans

6. Understanding Consumer Responses to Product Risk Information
Vol 70, No 1, January 2006, 79–91
Anthony D. Cox, Dena Cox, & Gregory Zimet

AWARD ANNOUNCEMENTS

7 Harold H. Maynard Award
Vol. 70, No. 2, April 2006, iv

8. Marketing Science Institute/H. Paul Root Award
Vol. 70, No. 2, April 2006, iv

9. Sheth Foundation/Journal of Marketing Award
Vol 70, No. 2, April 2006, iii

BRAND MANAGEMENT

10. Achieving Marketing Objectives Through Social Sponsorships
Vol. 70, No. 4, October 2006, 154–69
Carolyn J. Simmons & Karen L. Becker-Olsen

11. Brand Dilution: When Do New Brands Hurt Existing Brands?
Vol. 70, No. 2, April 2006, 52–66
Chris Pullig, Carolyn J Simmons, & Richard G Netemeyer

12. Brands as Signals: A Cross-Country Validation Study
Vol 70, No. 1, January 2006, 34–49
Tulin Erdem, Joffre Swait, & Ana Valenzuela

13. Drivers of Brand Extension Success
Vol 70, No. 2, April 2006, 18–34
Franziska Volckner & Henrik Sattler

14. Emotional Branding and the Strategic Value of the Doppelgänger Brand Image
Vol. 70, No. 1, January 2006, 50–64
Craig J. Thompson, Aric Rindfleisch, & Zeynep Arsel

15. The Evolution of Loyalty Intentions
Vol 70, No. 2, April 2006, 122–32
Michael D. Johnson, Andreas Herrmann, & Frank Huber

16. Human Brands: Investigating Antecedents to Consumers' Strong Attachments to Celebrities
Vol. 70, No. 3, July 2006, 104–119
Matthew Thomson

17. Ratios in Proportion: What Should the Shape of the Package Be?
Vol. 70, No. 2, April 2006, 95–107
Priya Raghubir & Eric A. Greenleaf

BUSINESS-TO-BUSINESS MARKETING

18. Friends, Businesspeople, and Relationship Roles: A Conceptual Framework and a Research Agenda
Vol. 70, No. 3, July 2006, 90–103
Jan B Heide & Kenneth H. Wathne

19. Influence Tactics for Effective Adaptive Selling
Vol. 70, No. 4, October 2006, 103–117
Richard G. McFarland, Goutam N. Challagalla, & Tasadduq A. Shervani

20. Market Pioneer and Early Follower Survival Risks: A Contingency Analysis of Really New Versus Incrementally New Product-Markets
Vol 70, No. 1, January 2006, 15–33
Sungwook Min, Manohar U Kalwani, & William T Robinson

21. Value-Based Differentiation in Business Relationships: Gaining and Sustaining Key Supplier Status
Vol. 70, No. 1, January 2006, 119–36
Wolfgang Ulaga & Andreas Eggert

CONSUMER BEHAVIOR

22. Brands as Signals A Cross-Country Validation Study
Vol 70, No. 1, January 2006, 34–49
Tülin Erdem, Joffre Swait, & Ana Valenzuela

© 2006, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

23. Broadening the Scope of Reference Price Advertising Research: A Field Study of Consumer Shopping Involvement
Vol. 70, No. 4, October 2006, 185–204
Daniel J. Howard & Roger A. Kerin

24. Consumer Contamination: How Consumers React to Products Touched by Others
Vol. 70, No. 2, April 2006, 81–94
Jennifer J. Argo, Darren W. Dahl, & Andrea C. Morales

25. Converting Web Site Visitors into Buyers: How Web Site Investment Increases Consumer Trusting Beliefs and Online Purchase Intentions
Vol. 70, No. 2, April 2006, 133–48
Ann E. Schlosser, Tiffany Barnett White, & Susan M. Lloyd

26. Crossover Dreams: Consumer Responses to Ethnic-Oriented Products
Vol. 70, No. 2, April 2006, 35–51
Sonya A. Grier, Anne M. Brumbaugh, & Corliss G. Thornton

27. Emotional Branding and the Strategic Value of the Doppelganger Brand Image
Vol. 70, No. 1, January 2006, 50–64
Craig J. Thompson, Aric Rindfleisch, & Zeynep Arsel

28. The Evolution of Loyalty Intentions
Vol. 70, No. 2, April 2006, 122–32
Michael D. Johnson, Andreas Herrmann, & Frank Huber

29. From Fear to Loathing? How Emotion Influences the Evaluation and Early Use of Innovations
Vol. 70, No. 3, July 2006, 44–57
Stacy L. Wood & C. Page Moreau

30. How Biased Household Inventory Estimates Distort Shopping and Storage Decisions
Vol. 70, No. 4, October 2006, 118–35
Pierre Chandon & Brian Wansink

31. How Country Characteristics Affect the Perceived Value of Web Sites
Vol. 70, No. 3, July 2006, 136–50
Jan-Benedict E.M. Steenkamp & Inge Geyskens

32. Human Brands: Investigating Antecedents to Consumers' Strong Attachments to Celebrities
Vol. 70, No. 3, July 2006, 104–119
Matthew Thomson

33. The Influence of Avatars on Online Consumer Shopping Behavior
Vol. 70, No. 4, October 2006, 19–36
Martin Holzwarth, Chris Janiszewski, & Marcus M. Neumann

34. Ratios in Proportion: What Should the Shape of the Package Be?
Vol. 70, No. 2, April 2006, 95–107
Priya Raghubir & Eric A. Greenleaf

35. The Role of Cognition and Affect in the Formation of Customer Satisfaction: A Dynamic Perspective
Vol. 70, No. 3, July 2006, 21–31
Christian Homburg, Nicole Koschate, & Wayne D. Hoyer

36. Scents and Sensibility: When Do (In)Congruent Ambient Scents Influence Product Evaluations?
Vol. 70, No. 3, July 2006, 32–43
Anick Bosmans

37. Understanding Consumer Responses to Product Risk Information
Vol. 70, No. 1, January 2006, 79–91
Anthony D. Cox, Dena Cox, & Gregory Zimet

38. The Unhealthy = Tasty Intuition and Its Effects on Taste Inferences, Enjoyment, and Choice of Food Products
Vol. 70, No. 4, October 2006, 170–84
Rajagopal Raghunathan, Rebecca Walker Naylor, & Wayne D. Hoyer

39. Upgrades and New Purchases
Vol. 70, No. 4, October 2006, 92–102
Erica Mina Okada

40. When Should a Retailer Create an Exciting Store Environment?
Vol. 70, No. 1, January 2006, 107–118
Velitchka D. Kaltcheva & Barton A. Weitz

CUSTOMER SATISFACTION

41. Are All Smiles Created Equal? How Emotional Contagion and Emotional Labor Affect Service Relationships
Vol. 70, No. 3, July 2006, 58–73
Thorsten Hennig-Thurau, Markus Groth, Michael Paul, & Dwayne D. Gremler

42. Corporate Social Responsibility, Customer Satisfaction, and Market Value
Vol. 70, No. 4, October 2006, 1–18
Xueming Luo & C.B. Bhattacharya

43. Customer Satisfaction and Stock Prices: High Returns, Low Risk
Vol. 70, No. 1, January 2006, 3–14
Claes Fornell, Sunil Mithas, Forrest V. Morgeson III, & M.S. Krishnan

44. From Fear to Loathing? How Emotion Influences the Evaluation and Early Use of Innovations
Vol. 70, No. 3, July 2006, 44–57
Stacy L. Wood & C. Page Moreau

45. The Role of Cognition and Affect in the Formation of Customer Satisfaction: A Dynamic Perspective
Vol. 70, No. 3, July 2006, 21–31
Christian Homburg, Nicole Koschate, & Wayne D. Hoyer

46. When Should a Retailer Create an Exciting Store Environment?
Vol. 70, No. 1, January 2006, 107–118
Velitchka D. Kaltcheva & Barton A. Weitz

DECISION MAKING

47. Brand Dilution: When Do New Brands Hurt Existing Brands?
Vol. 70, No. 2, April 2006, 52–66
Chris Pullig, Carolyn J. Simmons, & Richard G. Netemeyer

48. Does the Internet Promote Better Consumer Decisions? The Case of Name-Your-Own-Price Auctions
Vol. 70, No. 1, January 2006, 65–78
Martin Spann & Gerard J. Tellis

49. Friends, Businesspeople, and Relationship Roles: A Conceptual Framework and a Research Agenda
Vol. 70, No. 3, July 2006, 90–103
Jan B. Heide & Kenneth H. Wathne

50. How Biased Household Inventory Estimates Distort Shopping and Storage Decisions
Vol. 70, No. 4, October 2006, 118–35
Pierre Chandon & Brian Wansink

51. Stuck in the Past: Why Managers Persist with New Product Failures
Vol. 70, No. 2, April 2006, 108–121
Eyal Biyalogorsky, William Boulding, & Richard Staelin

52. Understanding Consumer Responses to Product Risk Information
Vol. 70, No. 1, January 2006, 79–91
Anthony D. Cox, Dena Cox, & Gregory Zimet

53. The Unhealthy = Tasty Intuition and Its Effects on Taste Inferences, Enjoyment, and Choice of Food Products
Vol. 70, No 4, October 2006, 170–84
Rajagopal Raghunathan, Rebecca Walker Naylor, & Wayne D. Hoyer

54. Upgrades and New Purchases
Vol. 70, No. 4, October 2006, 92–102
Erica Mina Okada

55. Valuing the Real Option of Abandoning Unprofitable Customers When Calculating Customer Lifetime Value
Vol. 70, No 3, July 2006, 5–20
Michael Haenlein, Andreas M. Kaplan, & Detlef Schoder

DIRECT MARKETING

56. It Just Feels Good: Customers' Affective Response to Touch and Its Influence on Persuasion
Vol. 70, No. 4, October 2006, 56–69
Joann Peck & Jennifer Wiggins

57. Valuing the Real Option of Abandoning Unprofitable Customers When Calculating Customer Lifetime Value
Vol. 70, No. 3, July 2006, 5–20
Michael Haenlein, Andreas M. Kaplan, & Detlef Schoder

EDITORIALS

58. From the Editor
Vol. 70, No. 1, January 2006, 1–2
Roland T. Rust

59. From the Editor: The Maturation of Marketing as an Academic Discipline
Vol 70, No. 3, July 2006, 1–2
Roland T. Rust

60. From the Web Editor. Improvements to the *Journal of Marketing* Web Site
Vol. 70, No. 3, July 2006, 3–4
Siva K. Balasubramanian

ELECTRONIC COMMERCE

61. Converting Web Site Visitors into Buyers: How Web Site Investment Increases Consumer Trusting Beliefs and Online Purchase Intentions
Vol 70, No. 2, April 2006, 133–48
Ann E. Schlosser, Tiffany Barnett White, & Susan M. Lloyd

62. Does the Internet Promote Better Consumer Decisions? The Case of Name-Your-Own-Price Auctions
Vol. 70, No. 1, January 2006, 65–78
Martin Spann & Gerard J. Tellis

63. How Country Characteristics Affect the Perceived Value of Web Sites
Vol. 70, No. 3, July 2006, 136–50
Jan-Benedict E M. Steenkamp & Inge Geyskens

64. The Influence of Avatars on Online Consumer Shopping Behavior
Vol. 70, No. 4, October 2006, 19–36
Martin Holzwarth, Chris Janiszewski, & Marcus M. Neumann

ETHICS AND SOCIAL RESPONSIBILITY

65. Achieving Marketing Objectives Through Social Sponsorships
Vol. 70, No. 4, October 2006, 154–69
Carolyn J. Simmons & Karen L. Becker-Olsen

66 Corporate Social Responsibility, Customer Satisfaction, and Market Value
Vol. 70, No. 4, October 2006, 1–18
Xueming Luo & C.B. Bhattacharya

67. The Unhealthy = Tasty Intuition and Its Effects on Taste Inferences, Enjoyment, and Choice of Food Products
Vol 70, No 4, October 2006, 170–84
Rajagopal Raghunathan, Rebecca Walker Naylor, & Wayne D Hoyer

GRAY MARKETS

68. How Does Enforcement Deter Gray Market Incidence?
Vol 70, No. 1, January 2006, 92–106
Kersi D. Antia, Mark E. Bergen, Shantanu Dutta, & Robert J. Fisher

INTERNATIONAL MARKETING

69. Brands as Signals. A Cross-Country Validation Study
Vol 70, No. 1, January 2006, 34–49
Tulin Erdem, Joffre Swait, & Ana Valenzuela

70. How Country Characteristics Affect the Perceived Value of Web Sites
Vol. 70, No 3, July 2006, 136–50
Jan-Benedict E M. Steenkamp & Inge Geyskens

LEGAL AND POLITICAL ISSUES

71. Brand Dilution: When Do New Brands Hurt Existing Brands?
Vol. 70, No. 2, April 2006, 52–66
Chris Pullig, Carolyn J. Simmons, & Richard G. Netemeyer

LOGISTICS/PHYSICAL DISTRIBUTION

72. Dual Distribution and Intangible Firm Value Franchising in Restaurant Chains
Vol. 70, No. 3, July 2006, 120–35
Raji Srinivasan

73. Friends, Businesspeople, and Relationship Roles: A Conceptual Framework and a Research Agenda
Vol. 70, No. 3, July 2006, 90–103
Jan B. Heide & Kenneth H. Wathne

MARKET OR INDUSTRY ANALYSIS

74. The Emergence of Dominant Designs
Vol. 70, No. 2, April 2006, 1–17
Raji Srinivasan, Gary L Lilien, & Arvind Rangaswamy

75 Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue
Vol. 70, No 3, July 2006, 74–89
Yong Liu

MARKET ORIENTATION

76 Creating a Market Orientation: A Longitudinal, Multifirm, Grounded Analysis of Cultural Transformation
Vol. 70, No. 4, October 2006, 37–55
Gary F. Gebhardt, Gregory S. Carpenter, & John F. Sherry Jr.

77. Customer Satisfaction and Stock Prices: High Returns, Low Risk
Vol. 70, No. 1, January 2006, 3–14
Claes Fornell, Sunil Mithas, Forrest V Morgeson III, & M.S. Krishnan

MARKETING CHANNELS

78 Dual Distribution and Intangible Firm Value: Franchising in Restaurant Chains
Vol. 70, No. 3, July 2006, 120–35
Raji Srinivasan

79. How Does Enforcement Deter Gray Market Incidence?
Vol 70, No. 1, January 2006, 92–106
Kersi D. Antia, Mark E. Bergen, Shantanu Dutta, &
Robert J. Fisher

MARKETING CONTROL

80 Marketing's Credibility: A Longitudinal Investigation of Marketing Communication Productivity and Shareholder Value
Vol 70, No. 4, October 2006, 70–91
Xueming Luo & Naveen Donthu

MARKETING IMPLEMENTATION

81. Factors Influencing the Effectiveness of Relationship Marketing. A Meta-Analysis
Vol. 70, No 4, October 2006, 136–53
Robert W. Palmatier, Rajiv P. Dant, Dhruv Grewal, &
Kenneth R Evans

ORGANIZATIONAL BUYING

82. Influence Tactics for Effective Adaptive Selling
Vol. 70, No 4, October 2006, 103–117
Richard G McFarland, Goutam N. Challagalla, &
Tasadduq A Shervani

83. Value-Based Differentiation in Business Relationships: Gaining and Sustaining Key Supplier Status
Vol. 70, No. 1, January 2006, 119–36
Wolfgang Ulaga & Andreas Eggert

ORGANIZATIONAL LEARNING

84. Creating a Market Orientation: A Longitudinal, Multifirm, Grounded Analysis of Cultural Transformation
Vol. 70, No. 4, October 2006, 37–55
Gary F Gebhardt, Gregory S. Carpenter, &
John F Sherry Jr.

85. Cross-Functional "Coopetition": The Simultaneous Role of Cooperation and Competition Within Firms
Vol. 70, No. 2, April 2006, 67–80
Xueming Luo, Rebecca J Slotegraaf, & Xing Pan

86 Stuck in the Past: Why Managers Persist with New Product Failures
Vol. 70, No. 2, April 2006, 108–121
Eyal Biyalogorsky, William Boulding, & Richard Staelin

ORGANIZATIONAL STRUCTURE

87. Cross-Functional "Coopetition": The Simultaneous Role of Cooperation and Competition Within Firms
Vol. 70, No. 2, April 2006, 67–80
Xueming Luo, Rebecca J Slotegraaf, & Xing Pan

88. Job Satisfaction, Job Performance, and Effort. A Reexamination Using Agency Theory
Vol 70, No 1, January 2006, 137–50
Markus Christen, Ganesh Iyer, & David Soberman

PERSONAL SELLING

89. The Influence of Avatars on Online Consumer Shopping Behavior
Vol 70, No. 4, October 2006, 19–36
Martin Holzwarth, Chris Janiszewski, &
Marcus M. Neumann

90. Influence Tactics for Effective Adaptive Selling
Vol. 70, No. 4, October 2006, 103–117
Richard G McFarland, Goutam N. Challagalla, &
Tasadduq A. Shervani

PRICING

91. Broadening the Scope of Reference Price Advertising Research: A Field Study of Consumer Shopping Involvement
Vol 70, No. 4, October 2006, 185–204
Daniel J. Howard & Roger A Kerin

92. Does the Internet Promote Better Consumer Decisions? The Case of Name-Your-Own-Price Auctions
Vol 70, No. 1, January 2006, 65–78
Martin Spann & Gerard J. Tellis

PRODUCT DEVELOPMENT

93. The Emergence of Dominant Designs
Vol. 70, No. 2, April 2006, 1–17
Raji Srinivasan, Gary L. Lilien, & Arvind Rangaswamy

94. From Fear to Loathing? How Emotion Influences the Evaluation and Early Use of Innovations
Vol. 70, No 3, July 2006, 44–57
Stacy L Wood & C. Page Moreau

95 Market Pioneer and Early Follower Survival Risks: A Contingency Analysis of Really New Versus Incrementally New Product-Markets
Vol 70, No. 1, January 2006, 15–33
Sungwook Min, Manohar U. Kalwani, &
William T Robinson

96. Ratios in Proportion: What Should the Shape of the Package Be?
Vol. 70, No. 2, April 2006, 95–107
Priya Raghubir & Eric A. Greenleaf

97. The Role of Cognition and Affect in the Formation of Customer Satisfaction: A Dynamic Perspective
Vol. 70, No. 3, July 2006, 21–31
Christian Homburg, Nicole Koschate, & Wayne D Hoyer

98 Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue
Vol. 70, No 3, July 2006, 74–89
Yong Liu

PRODUCT MANAGEMENT

99. Drivers of Brand Extension Success
Vol. 70, No. 2, April 2006, 18–34
Franziska Völckner & Henrik Sattler

100. The Emergence of Dominant Designs
Vol. 70, No. 2, April 2006, 1–17
Raji Srinivasan, Gary L Lilien, & Arvind Rangaswamy

101 Stuck in the Past: Why Managers Persist with New Product Failures
Vol 70, No. 2, April 2006, 108–121
Eyal Biyalogorsky, William Boulding, & Richard Staelin

RELATIONSHIP MARKETING

102. Factors Influencing the Effectiveness of Relationship Marketing: A Meta-Analysis
Vol. 70, No. 4, October 2006, 136–53
Robert W Palmatier, Rajiv P. Dant, Dhruv Grewal, &
Kenneth R Evans

RETAILING

103. Consumer Contamination: How Consumers React to Products Touched by Others
Vol. 70, No. 2, April 2006, 81–94
Jennifer J Argo, Darren W. Dahl, & Andrea C. Morales

104. Converting Web Site Visitors into Buyers: How Web Site Investment Increases Consumer Trusting Beliefs and Online Purchase Intentions
Vol. 70, No. 2, April 2006, 133–48
Ann E. Schlosser, Tiffany Barnett White, & Susan M. Lloyd

105. How Biased Household Inventory Estimates Distort Shopping and Storage Decisions
Vol. 70, No. 4, October 2006, 118–35
Pierre Chandon & Brian Wansink

106. It Just Feels Good: Customers' Affective Response to Touch and Its Influence on Persuasion
Vol. 70, No. 4, October 2006, 56–69
Joann Peck & Jennifer Wiggins

107. Scents and Sensibility: When Do (In)Congruent Ambient Scents Influence Product Evaluations?
Vol. 70, No. 3, July 2006, 32–43
Anick Bosmans

108. When Should a Retailer Create an Exciting Store Environment?
Vol. 70, No. 1, January 2006, 107–118
Velitchka D. Kaltcheva & Barton A. Weitz

SALES FORCE MANAGEMENT

109. Factors Influencing the Effectiveness of Relationship Marketing: A Meta-Analysis
Vol. 70, No. 4, October 2006, 136–53
Robert W. Palmatier, Rajiv P Dant, Dhruv Grewal, & Kenneth R. Evans

110. Job Satisfaction, Job Performance, and Effort: A Reexamination Using Agency Theory
Vol. 70, No. 1, January 2006, 137–50
Markus Christen, Ganesh Iyer, & David Soberman

SALES FORECASTING

111. Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue
Vol. 70, No. 3, July 2006, 74–89
Yong Liu

SEGMENTATION

112. Crossover Dreams: Consumer Responses to Ethnic-Oriented Products
Vol. 70, No. 2, April 2006, 35–51
Sonya A. Grier, Anne M. Brumbaugh, & Corliss G. Thornton

113. Upgrades and New Purchases
Vol. 70, No. 4, October 2006, 92–102
Erica Mina Okada

114. Valuing the Real Option of Abandoning Unprofitable Customers When Calculating Customer Lifetime Value
Vol. 70, No. 3, July 2006, 5–20
Michael Haenlein, Andreas M. Kaplan, & Detlef Schoder

SERVICE QUALITY

115. Customer Satisfaction and Stock Prices: High Returns, Low Risk
Vol. 70, No. 1, January 2006, 3–14
Claes Fornell, Sunil Mithas, Forrest V. Morgeson III, & M.S. Krishnan

SERVICES MARKETING

116. Are All Smiles Created Equal? How Emotional Contagion and Emotional Labor Affect Service Relationships
Vol 70, No. 3, July 2006, 58–73
Thorsten Hennig-Thurau, Markus Groth, Michael Paul, & Dwayne D. Gremler

117. Emotional Branding and the Strategic Value of the Doppelgänger Brand Image
Vol 70, No. 1, January 2006, 50–64
Craig J. Thompson, Aric Rindfleisch, & Zeynep Arsel

118. The Evolution of Loyalty Intentions
Vol. 70, No. 2, April 2006, 122–32
Michael D. Johnson, Andreas Herrmann, & Frank Huber

SPONSORSHIP

119. Achieving Marketing Objectives Through Social Sponsorships
Vol. 70, No. 4, October 2006, 154–69
Carolyn J. Simmons & Karen L. Becker-Olsen

STRATEGIC MARKETING

120. Corporate Social Responsibility, Customer Satisfaction, and Market Value
Vol. 70, No. 4, October 2006, 1–18
Xueming Luo & C.B. Bhattacharya

121. Creating a Market Orientation: A Longitudinal, Multifirm, Grounded Analysis of Cultural Transformation
Vol. 70, No. 4, October 2006, 37–55
Gary F Gebhardt, Gregory S. Carpenter, & John F. Sherry Jr.

122. Cross-Functional "Coopetition": The Simultaneous Role of Cooperation and Competition Within Firms
Vol. 70, No. 2, April 2006, 67–80
Xueming Luo, Rebecca J. Slotegraaf, & Xing Pan

123. Dual Distribution and Intangible Firm Value: Franchising in Restaurant Chains
Vol. 70, No. 3, July 2006, 120–35
Raji Srinivasan

124. Market Pioneer and Early Follower Survival Risks: A Contingency Analysis of Really New Versus Incrementally New Product-Markets
Vol. 70, No. 1, January 2006, 15–33
Sungwook Min, Manohar U. Kalwani, & William T. Robinson

125. Marketing's Credibility: A Longitudinal Investigation of Marketing Communication Productivity and Shareholder Value
Vol. 70, No. 4, October 2006, 70–91
Xueming Luo & Naveen Donthu

126 Value-Based Differentiation in Business Relationships: Gaining and Sustaining Key Supplier Status
Vol 70, No. 1, January 2006, 119–36
Wolfgang Ulaga & Andreas Eggert

AUTHOR INDEX

	Item Number
Antia, Kersi D.	68, 79
Argo, Jennifer J.	24, 103
Arsel, Zeynep	14, 27, 117
Balasubramanian, Siva K	60
Becker-Olsen, Karen L.	10, 65, 119
Bergen, Mark E.	68, 79
Bhattacharya, C B.	42, 66, 120
Biyalogorsky, Eyal	51, 86, 101
Bosmans, Anick	5, 36, 107
Boulding, William	51, 86, 101
Brumbaugh, Anne M.	2, 26, 112
Carpenter, Gregory S.	76, 84, 121
Challagalla, Goutam N.	19, 82, 90
Chandon, Pierre	30, 50, 105
Christen, Markus	88, 110
Cox, Anthony D.	6, 37, 52
Cox, Dena	6, 37, 52
Dahl, Darren W.	24, 103
Dant, Rajiv P.	81, 102, 109
Donthu, Naveen	4, 80, 125
Dutta, Shantanu	68, 79
Eggert, Andreas	21, 83, 126
Erdem, Tulin	12, 22, 69
Evans, Kenneth R.	81, 102, 109
Fisher, Robert J.	68, 79
Fornell, Claes	43, 77, 115
Gebhardt, Gary F.	76, 84, 121
Geyskens, Inge	31, 63, 70
Greenleaf, Eric A.	17, 34, 96
Gremler, Dwayne D.	41, 116
Grewal, Dhruv	81, 102, 109
Grier, Sonya A.	2, 26, 112
Groth, Markus	41, 116
Haenlein, Michael	55, 57, 114
Heide, Jan B.	18, 49, 73
Hennig-Thurau, Thorsten	41, 116
Herrmann, Andreas	15, 28, 118
Holzwarth, Martin	33, 64, 89
Homburg, Christian	35, 45, 97
Howard, Daniel J.	1, 23, 91
Hoyer, Wayne D.	35, 38, 45, 53, 67, 97
Huber, Frank	15, 28, 118
Iyer, Ganesh	88, 110
Janiszewski, Chris	33, 64, 89
Johnson, Michael D.	15, 28, 118
Kaltcheva, Velitchka D.	40, 46, 108
Kalwani, Manohar U.	20, 95, 124
Kaplan, Andreas M.	55, 57, 114
Kerin, Roger A	1, 23, 91
Koschate, Nicole	35, 45, 97
Krishnan, M.S.	43, 77, 115
Lilien, Gary L.	74, 93, 100
Liu, Yong	75, 98, 111
Lloyd, Susan M.	25, 61, 104
Luo, Xueming	4, 42, 66, 80, 85, 87, 120, 122, 125
McFarland, Richard G.	19, 82, 90
Min, Sungwook	20, 95, 124
Mithas, Sunil	43, 77, 115
Morales, Andrea C.	24, 103
Moreau, C Page	29, 44, 94
Morgeson, Forrest V., III	
Naylor, Rebecca Walker	38, 53, 67
Netemeyer, Richard G	11, 47, 71
Neumann, Marcus M.	33, 64, 89
Okada, Erica Mina	39, 54, 113
Palmatier, Robert W	81, 102, 109
Pan, Xing	85, 87, 122
Paul, Michael	41, 116
Peck, Joann	3, 56, 106
Pullig, Chris	11, 47, 71
Raghubir, Priya	17, 34, 96
Raghunathan, Rajagopal	38, 53, 67
Rangaswamy, Arvind	74, 93, 100
Rindfleisch, Aric	14, 27, 117
Robinson, William T.	20, 95, 124
Rust, Roland T.	58, 59
Sattler, Henrik	13, 99
Schlosser, Ann E.	25, 61, 104
Schoder, Detlef	55, 57, 114
Sherry, John F., Jr.	
Shervani, Tasadduq A.	19, 82, 90
Simmons, Carolyn J.	10, 11, 47, 65, 71, 119
Slotegraaf, Rebecca J.	85, 87, 122
Soberman, David	88, 110
Spann, Martin	48, 62, 92
Srinivasan, Raji	72, 74, 78, 93, 100, 123
Staelin, Richard	51, 86, 101
Steenkamp, Jan-Benedict E.M.	31, 63, 70
Swait, Joffre	12, 22, 69
Tellis, Gerard J	48, 62, 92
Thompson, Craig J.	14, 27, 117
Thomson, Matthew	16, 32
Thornton, Corliss G.	2, 26, 112
Ulaga, Wolfgang	21, 83, 126
Valenzuela, Ana	12, 22, 69
Völckner, Franziska	13, 99
Wansink, Brian	30, 50, 105
Wathne, Kenneth H.	18, 49, 73
Weitz, Barton A.	40, 46, 108
White, Tiffany Barnett	25, 61, 104
Wiggins, Jennifer	3, 56, 106
Wood, Stacy L	29, 44, 94
Zimet, Gregory	6, 37, 52

TEXERE – The Intersection of Insight & Innovation

THOMSON

AMERICAN MARKETING ASSOCIATION

The **American Marketing Association** is proud to have a strategic alliance with Texere, a product line of Thomson Professional Business, to deliver a new generation of professional marketing resources to help you better meet the demands of today's challenging marketing arena. Our vision for the new AMA Bookstore is to bring you a full library of the most cutting-edge resources for marketing professionals filled with practical marketing solutions and insight into innovative marketing strategies.

NEW Books from the AMA Bookstore

Customer Message Management
It is no longer what you sell, but how you sell it—meaning it is now vital to bridge the gap between sales and marketing. *Customer Message Management* does just that, eradicating the disconnect that for years has plagued the two departments
Tim Riesterer
0-324-31316-0 — $34.95 / AMA Members $27.95

Competing for Customers and Capital
Competing for Customers and Capital illustrates a cause-and-effect model of relationships between marketing and finance based on a common language, economic theory, and financial accounting data.
Dr. Victor J. Cook Jr.
0-324-40597-9 — $39.95 / AMA Members $31 95

Managing Risks for Corporate Integrity: How to Survive an Ethical Misconduct Disaster
This text offers an inside look at recent ethical disasters at prominent American corporations as it teaches readers how to recognize, avoid, discover, answer, and recover (RADAR) from ethical mishaps in their own companies.
Lynn Brewer, Robert Chandler, and O.C. Ferrell
0-324-20351-9 — $27.95 / AMA Members $22.50

The Marketer's Guide to Public Relations in the 21st Century
Packed with riveting examples from the most sophisticated practitioners—including the success of products as diverse as Apple's iPod, Viagra, Botox, Toyota Prius, Sony Playstation 2, and the Harry Potter books—this new book illustrates how MPR helps companies gain a competitive edge in today's cluttered marketplace
Thomas L. Harris and Patricia T. Whalen
0-324-31210-5 — $39.95 / AMA Members $31 95

The White Paper Marketing Handbook
As consumers become more resistant to direct sales appeals, white papers—promotional pieces in the guise of informational articles or reports—are increasingly alluring. This exciting new book is the first to tell marketers how to create effective white papers using a variety of media.
Robert W. Bly
0-324-30082-4 — $39.95 / AMA Members $31 95

Farce to Force: Building Profitable E-Commerce Strategies
Though it's a powerful driving force for commerce, early misuse of business principles can lead to bad business practices—as well as perceptions that the Internet is not a viable sales channel *Farce to Force* sets the record straight with techniques and examples of successful e-commerce marketing
Sarah McCue
0-538-72677-6 — $27 95 / AMA Members $22.50

Managing Global Accounts: Nine Critical Factors for a World-Class Program
Drawing on their experiences, the authors detail nine critical success factors for designing and executing a world-class global account management program.
Noel Capon, Dave Potter, and Fred Schindler
0-324-40076-4 — $29.95 / AMA Members $23.95

Strategic Sales Leadership: BREAKthrough Thinking for BREAKthrough Results
While there are numerous cookbook approaches to sales management, very little is available on training executive sales managers how to think and manage their sales teams strategically Until now. This book equips sales executives with the tools to develop an effective action plan to improve sales organization results.
The Sales Educators
1-58799-203-5 — $39.95 / AMA Members $31 95

Ten Key Customer Insights: Unlocking the Mind of the Market
While companies tout "Customer is King," the reality is declining market share, new product failure, and financial woes. This book helps readers develop marketing strategies and tactics to attract and retain more customers, grow market share, increase the productivity of their marketing efforts, and boost profitability.
Robert Schieffer
1-58799-206-X — $29.95 / AMA Members $23.95

Creating Do-it-Yourself Customers: How Great Customer Experiences Build Great Companies
Self-checkout, self-customization, and other co-production efforts are on the rise But it's a boon to business only if customers do their share. This book helps readers recognize situations where high-tech or human innovation can bolster customer involvement
Peter Honebein and Roy Cammarano
0-324-31154-0 — $29.95 / AMA Members $23 95

To Order: 1-800-842-3636 http://www.marketingpower.com/amabooks

You have to see it to believe it.

It's a marketing event like no other.

A journey of **unparalleled discovery** offering ground-breaking solutions to today's most pressing marketing challenges.

And you're invited.

Join the **American Marketing Association** and step outside the realm of the typical marketing conference:

- **> EXPLORE** the new frontier of marketing for fresh insights and never-before-seen research on connecting with customers.
- **> TRANSFORM** your marketing organization and change mindsets with innovative tools and techniques that drive actionable results enterprise wide.
- **> CONNECT** with peers in a podium-free environment focused on collaboration and interaction.
- **> EXPERIENCE** over two days of total marketing immersion with industry leaders and rising stars on an unprecedented journey of knowledge and inspiration.

M planet™
AMERICAN MARKETING ASSOCIATION

Start your journey today: **Mplanet2006.com**

November 29–December 1, 2006
Walt Disney World Dolphin Resort